



Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes

Yuma Akbar^{1*}, Tri Sugiharto²

^{1,2}Sistem Informasi, St Ilmu Komputer Cipta Karya Informatika, Jakarta, Indonesia

Email Penulis Korespondensi: ¹yuma.pjj@gmail.com

Abstrak—Perkembangan teknologi di era sekarang terpaut pesat, terlebih beberapa akhir ini dan terbaru mengenai tentang kecerdasan buatan. Banyak beberapa raksasa teknologi mengembangkan kecerdasan buatan yang bertujuan untuk mempermudah pekerjaan manusia. Salah satunya adalah ChatGPT. ChatGPT dikembangkan oleh OpenAI, perusahaan riset kecerdasan buatan yang berbasis di San Francisco, California, Amerika Serikat, pada tahun 2020 dan diperkenalkan secara resmi kepada publik pada tahun 2021. ChatGPT dikembangkan berdasarkan model bahasa generatif terdahulu sebelumnya, yaitu GPT-3. Keberadaan ChatGPT menjadi bukti kemajuan teknologi chatbot saat ini. Penelitian yang saya lakukan mengusung tema Analisis Sentimen Pengguna Twitter di Indonesia terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes. Data penelitian bersumber dari pengguna Twitter, spesifik di Indonesia. Penelitian ini bertujuan untuk menganalisis pola penggunaan Twitter di Indonesia terhadap ChatGPT menggunakan algoritma C4.5 dan Naïve Bayes. Dengan menggunakan RapidMiner, metode C4.5 digunakan untuk mengidentifikasi pola penggunaan Twitter yang paling signifikan dalam interaksi dengan ChatGPT, sedangkan metode Naïve Bayes digunakan untuk mengklasifikasikan jenis interaksi yang paling umum terjadi antara pengguna Twitter dan ChatGPT. Perbandingan kedua algoritma mendapat hasil accuracy = 77.33%, precision = 100.00%, recall = 30.18%. Kesimpulan dari hasil tersebut adalah positif, artinya respon atau sentimen yang diambil dari sosial media twitter khususnya di Indonesia menunjukkan feedback baik terhadap ChatGPT.

Kata Kunci: ChatGPT, Twitter, Analisis Sentimen, C4.5, Naïve Bayes

Abstract—Technological developments in the current era are rapidly adrift, especially recently and the latest regarding artificial intelligence. Many technology giants are developing artificial intelligence which aims to make human work easier. One of them is ChatGPT. ChatGPT was developed by OpenAI, an artificial intelligence research company based in San Francisco, California, United States, in 2020 and officially introduced to the public in 2021. ChatGPT was developed based on the previous deepest generative language model, namely GPT-3. The existence of ChatGPT is evidence of the advancement of chatbot technology today. The research that I did carried out the theme of Sentiment Analysis of Twitter Users in Indonesia on ChatGPT Using the C4.5 Algorithm and Naïve Bayes. Research data comes from Twitter users, specifically in Indonesia. This study aims to analyze the pattern of Twitter usage in Indonesia towards ChatGPT using the C4.5 and Naïve Bayes algorithms. Using RapidMiner, the C4.5 method is used to identify the most significant patterns of Twitter usage in interactions with ChatGPT, while the Naïve Bayes method is used to classify the most common types of interactions that occur between Twitter and ChatGPT users. Comparison of the two algorithms results in accuracy = 77.33%, precision = 100.00%, recall = 30.18%. The conclusion from these results is positive, meaning that the response or sentiment taken from social media Twitter, especially in Indonesia, shows good feedback to ChatGPT.

Keywords: ChatGPT, Twitter, Sentiment Analysis, C4.5, Naïve Bayes

I. PENDAHULUAN

Penggunaan chatbot semakin populer dan berkembang pesat. Chatbot dapat membantu memudahkan interaksi antara pengguna dan perusahaan atau organisasi, seperti customer service (*Live Chat*). [1] Bahkan, di era sekarang kecerdasan buatan melalui chatbot menjadi salah satu andalan dalam pemecahan berbagai masalah. Salah satu jenis chatbot yang populer adalah ChatGPT, yang menggunakan teknologi NLP (*Natural Language Processing*) untuk memahami dan merespons pertanyaan dan permintaan pengguna. [2]

Penelitian ini, fokus pada pengguna Twitter di Indonesia karena Indonesia adalah salah satu negara dengan pengguna aktif Twitter terbesar di dunia. Oleh karena itu, dengan parameter tagar atau tanda pagar

#ChatGPT menjadi acuan untuk pengambilan data atau opini tentang respon warganet dalam penelitian ini. [3]

Belum ada penelitian yang secara khusus menganalisis sentimen pengguna Twitter di Indonesia terhadap ChatGPT. Penelitian ini bertujuan untuk mengisi kesenjangan tersebut dan memberikan pemahaman yang lebih dalam mengenai sentimen pengguna Twitter di Indonesia terhadap ChatGPT. Beberapa masalah yang perlu diidentifikasi, yaitu :

1. Apakah kualitas data yang digunakan, yakni data yang digunakan harus relevan dengan tema penelitian dan tidak mengandung spam.
2. Apakah mengelola data sentimen harus memperhatikan keamanan dan privasi pengguna sesuai kebijakan privasi twitter dan aturan hukum yang berlaku.



Berdasarkan identifikasi masalah tersebut, didapat rumusan masalah sebagai berikut :

1. Bagaimana cara mengumpulkan data tweet yang relevan untuk analisis sentimen pengguna Twitter di Indonesia terhadap ChatGPT?
2. Bagaimana perbandingan performa antara algoritma C4.5 dan Naive Bayes dalam mengklasifikasikan sentimen tweet pengguna Twitter di Indonesia terhadap ChatGPT?

Oleh karena itu, mengetahui apakah pengguna Twitter di Indonesia memiliki sentimen positif atau negatif terhadap adanya ChatGPT, serta mengevaluasi kedua algoritma dalam melakukan analisis sentimen pada media sosial.

Definisi dan Pengertian

1. Analisis Sentimen

Analisis sentimen adalah proses pengumpulan, pengolahan, dan evaluasi opini atau pendapat publik terhadap suatu topik atau produk tertentu.[4]

2. Artificial Intelligence

Kecerdasan buatan mencakup berbagai aspek seperti pemrosesan bahasa alami, visi komputer, pengambilan keputusan, pembelajaran mesin, dan robotika.[5]

3. ChatGPT

ChatGPT dikembangkan oleh OpenAI, perusahaan riset kecerdasan buatan yang berbasis di San Francisco, California, Amerika Serikat, pada tahun 2020 dan diperkenalkan secara resmi kepada publik pada tahun 2021. ChatGPT dikembangkan berdasarkan model bahasa generatif terdahulu sebelumnya, yaitu GPT-3.[6]

4. Crawling Data

Crawling adalah istilah dalam teknologi informasi yang merujuk pada proses pengumpulan data dari internet secara otomatis dengan menggunakan perangkat lunak khusus yang disebut web crawler atau spider.[7]

5. Flowchart

Flowchart adalah diagram atau gambar yang digunakan untuk merepresentasikan urutan langkah-langkah atau proses dalam suatu sistem atau kegiatan.[8]

II. METODE PENELITIAN

A. Metode

Metode yang digunakan pada penelitian adalah C4.5 dan Naive Bayes. metode C4.5 digunakan untuk mengidentifikasi pola penggunaan Twitter yang paling signifikan dalam interaksi dengan ChatGPT, sedangkan metode Naive Bayes digunakan untuk mengklasifikasikan jenis interaksi yang paling umum terjadi antara pengguna Twitter dan ChatGPT.

B. Data Penelitian

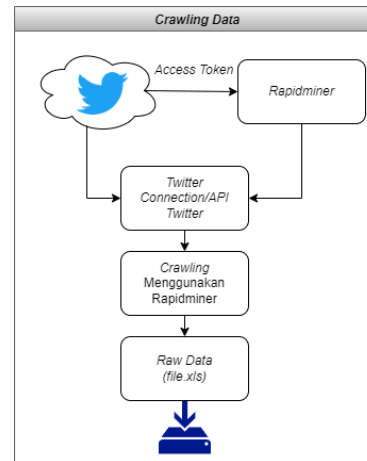
Data dalam penelitian ini bersifat OpenData atau publik. Bersumber dari opini (komentar) warganet Indonesia di sosial media Twitter mengenai ChatGPT. Jumlah data yang digunakan adalah 1607 tweets pada periode Maret – April 2023. Metode pendekatan yakni kuantitatif dengan mengambil dan mengolah data dari sosial media twitter menggunakan tools Rapidminer. Proses untuk mendapatkan data penelitian :

1. Studi Literatur

Studi pustaka dengan melakukan pengumpulan teori terkait dari berbagai sumber literasi yaitu jurnal, *e-book* dan artikel.[9]

2. Crawling Data

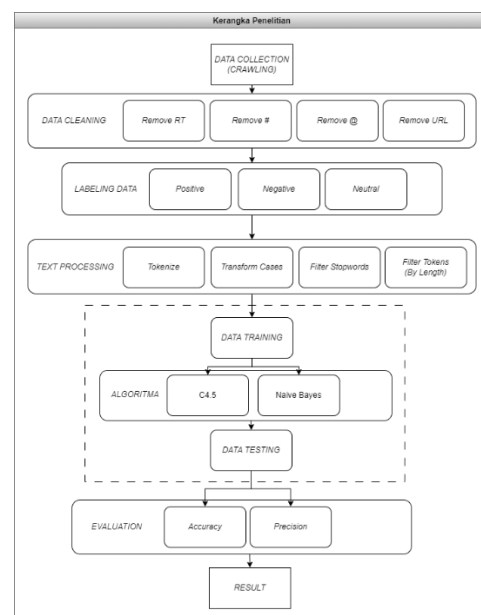
Mengumpulkan dan mengindeks data dari sosial media twitter menggunakan Rapidminer.[10]



Gambar 1. Crawling Data

C. Penerapan Metodologi

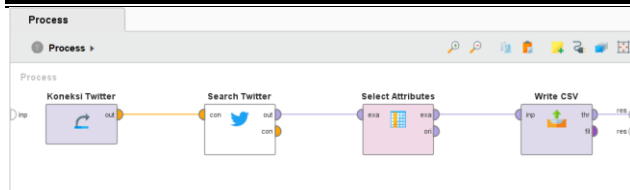
Penerapan metodologi pada penelitian ini dilakukan beberapa proses atau tahapan, diantaranya pengumpulan data, pembersihan data, *pre-processing* data, penerapan algoritma, evaluasi performa dan kesimpulan/hasil. Berikut adalah alur proses penerapan metodologi :



Gambar 2. Kerangka Penelitian

1. Pengumpulan Data

Pengumpulan data atau crawling menggunakan operator *search* twitter pada rapidminer. Membangun koneksi antara rapidminer dan API twitter dengan sistem tokenisasi untuk dapat akses.[11]



Gambar 3. Proses Crawling Data di Rapidminer

Row No.	Text
1500	1x cok 1k
1501	1x cok 1k
1502	1x cok 1k
1503	deteksi di chatGPT tidak
1504	1 atau kerdil yang chat
1505	1 FREE PDF PENGUNYONGAN
1506	1 ada proof NO RESPONSI ORNY
1507	https://t.co/1YMMW0WZ
1508	Kopikan tanggapan chatgpt plus. Worth it kan?
1509	1 25w app premium
1510	1 AU NEEDED & Double LYSN
1511	1 laptop games
1512	1 chatgpt, https://t.co/0u5Qf7mZS , telegram & studynote
1513	1 elvira samah
1514	yai om per https://t.co/1tC4k05iQy
1515	@elviratm mana lo da dapet nia H0370 jga lu? ChatGPT?
1516	@githny terkalahkan chatgpt!
1517	RT @AlkabarRBM: Saya bahkan minta ChatGPT untuk membuat saya link lagu tentang Blockchain (UAR RASA. Dst awal sampai akhir dibuat...

Gambar 4. Dataset

2. Data Cleaning

Proses ini merupakan menormalisasi atau menyortir tweets guna memberikan data relevan serta terhindar dari spam. Dilakukan penghapusan RT, @, # dan URL.[12]

Tabel 1. Data Cleaning

SEBELUM	SESUDAH
@selfolrps Tadi nya jabarin fotonya liat foto nya, sekarang biar wording nya lebih alus udah dapet basic nya di convert ke chatgpt ai jadi hasilnya lebih bagus. Yang penting kalimat basic dari foto yg mau di upchar udah ada nanti gampang deh peralus wording	Tadi nya jabarin fotonya liat foto nya, sekarang biar wording nya lebih alus udah dapet basic nya di convert ke chatgpt ai jadi hasilnya lebih bagus. Yang penting kalimat basic dari foto yg mau di upchar udah ada nanti gampang deh peralus wording
Waspada, Banyak Pria Lajang Tertarik Gunakan ChatGPT untuk Tipu Calon Pasangan Kencan. - #RiauPos #Kaspersky #aplikasiKencanOnline #chatGpt https://t.co/zQR0pWtJY	Waspada, Banyak Pria Lajang Tertarik Gunakan ChatGPT untuk Tipu Calon Pasangan Kencan.
RT @borryshasian: Iseng tanya ke ChatGPT, gimana cara survive di Indo, kalo ga bisa narik cash dan ga bisa pake m-banking.	Iseng tanya ke ChatGPT, gimana cara survive di Indo, kalo ga bisa narik cash dan ga bisa pake m-banking.

3. Pelabelan Data

Pelabelan data dilakukan secara manual pada Ms. Excel dengan menentukan klasifikasi sentimen dari tweets, bersifat positif, negatif atau netral. Sentimen netral berisi opini biasa. Sentimen positif berisi opini dan komentar yang mendukung tentang ChatGPT, sebaliknya untuk sentimen negatif berisi opini dan komentar yang tidak mendukung.[13]

Tabel 2. Pelabelan Data

Text	Sentimen
Tadi nya jabarin fotonya liat foto nya, sekarang biar wording nya lebih alus udah dapet basic nya di convert ke chatgpt ai jadi hasilnya lebih bagus. Yang penting kalimat basic dari foto yg mau di upchar udah ada nanti gampang deh peralus wording	Positive
Waspada, Banyak Pria Lajang Tertarik Gunakan ChatGPT untuk Tipu Calon Pasangan Kencan.	Negative

4. Pre-processing Data

Beberapa tahapan dalam pemrosesan dataset pada penelitian ini :

1) Tokenize

Pada proses ini tweets akan dipecah menjadi per kata, maksudnya memisah kan satu kata dengan kata lain. Setiap kata akan dijadikan satu atribut tersendiri. Menghasilkan data (TF-IDF). TF-IDF adalah suatu teknik untuk menghitung berapa kali suatu kata muncul dalam sebuah dokumen.

2) Transform Cases

Proses ini menyamaratakan ukuran font agar seragam. Misal mengubah font ke huruf besar seluruhnya atau mengubah ke huruf kecil seluruhnya.

3) Filter Stopwords

Di proses ini, menghapus/membuang kata-kata tidak penting. Seperti yang, dan, di, adalah dll.

4) Filter Tokens (By Length)

Tahapan ini menghapus/membuang kata-kata yang mengandung sedikit huruf atau menghapus/membuang kata-kata yang mengandung banyak huruf.

5) Wordcloud

Tahap tampilan bobot kata yang dominan muncul, divisualisasikan dalam bentuk wordcloud.

5. Penerapan Algoritma

Pada penelitian ini menerapkan 2 algoritma, yaitu :

1) Algoritma C4.5

Digunakan algoritma C4.5 untuk membentuk pohon keputusan dengan istilah-istilah yang merupakan fitur utama dari data teks sebagai parameter.[14] Dalam proses memilih atribut(kata) sebagai akar, perlu menghitung entropy dan gain. Untuk menghitung gain dilakukan dengan rumus berikut :

$$Gain(A) = Entropy(S); \sum_{i=1}^n P_i \cdot \log_2(P_i)$$

Dimana :

S = Himpunan kasus

A = Atribut

n = Jumlah kasus dalam S



Selanjutnya untuk menghitung nilai entropy dengan rumus :

$$Entropy(S_1, S_2, \dots, S_n) = \sum_{i=1}^n P_i \cdot \log_2$$

Dengan :

S = Himpunan kasus

n = Jumlah Sampel

Pi = Proporsi kelas

2) Algoritma Naïve Bayes

Dalam algoritma Naïve bayes setiap dokumen diwakili oleh sepasang atribut “x1, x2, x3,...xn” dimana x1 adalah kata pertama, x2 adalah kata kedua dan seterusnya. Sedangkan V adalah kumpulan kategori Tweet(opini). Selama klasifikasi, algoritma mencari probabilitas tertinggi di antara semuanya.[15] Kategori Dokumen Teruji (VMAP) memiliki kesamaan adalah sebagai berikut :

$$P(V_j) = \frac{|docs\ j|}{|kosakata|}$$

Keterangan :

|docs j| = Jumlah dokumen dalam kategori j

|contoh| = Jumlah dokumen dari semua kategori

nk = Jumlah frekuensi kemunculan setiap kata

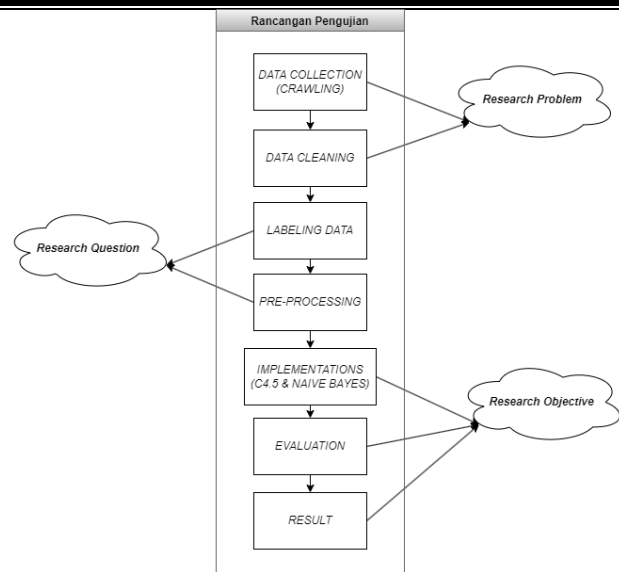
n = Jumlah frekuensi kemunculan kata dari setiap kategori

|kosakata| = Jumlah semua kata dari setiap kategori

D. Rancangan Pengujian

Rancangan pengujian menyeluruh dari penelitian ini bertujuan untuk menjawab identifikasi masalah dan rumusan masalah.

1. Pengumpulan data menggunakan teknik crawling menggunakan rapidminer.
2. Dataset dibersihkan atau dinormalisasi guna menghasilkan data(tweets) yang relevan dan terhindar dari spam.
3. Pelabelan data menggunakan Ms. Excel dengan menentukan klasifikasi positif, negatif atau netral.
4. *Pre-processing* data dengan 5 operator : *tokenize*, *transforms case*, *filter stopwords*, *filter tokens(by length)* dan *wordcloud*.
5. Penerapan algoritma C4.5 dan Naïve Bayes pada pola dataset, mengevaluasi performa antar algoritma dan menentukan hasil dari sentimen.



Gambar 5. Rancangan Pengujian

III. HASIL DAN PEMBAHASAN

A. Implementasi dan Pengujian

Tahap implementasi dan pengujian menggunakan beberapa metodologi sebagai berikut :

1. Pengumpulan Data

Pengumpulan data menggunakan teknik crawling, mengindeks opini atau komentar pengguna twitter. Setelah koneksi terbangun antara rapidminer dengan twitter lalu opini diindeks dengan output file.csv. Parameter pencarian = query : 'ChatGPT', result type : 'recent or popular', limit : '2000', language : 'id'. Tambah operators select attributes dengan parameter = type : 'include attributes', attribute filter types = 'one attribute', select attribute : 'Text', centang checkbox 'also apply to special attributes'. Terakhir tambah operator write xls, untuk mendapatkan data dalam bentuk file.csv. Didapatkan total 1607 tweets.

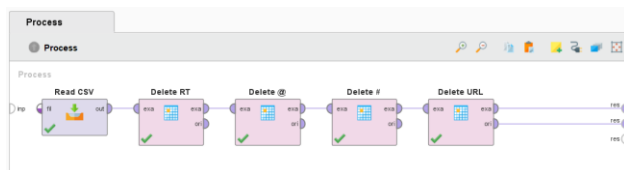
Row No.	Text
1590	Si cek GK
1591	7x cek GK
1592	10x cek GK
1593	deteksi AI chatGPT Skolek
1594	7 akan bunteri start skolek
1595	7 FREE PDF PANDUAN PENGUNJUAN
1596	7 ada proof NO HAKSIH CHAT
1597	https://t.co/1V00MM02
1598	Kepuasan tanggapan chatgpt plus. Worth it kah?
1599	7 20+ app premium
1600	7 AU NEEDS & Bubble LYNN
1601	7 laptop games
1602	7 chatgpt, https://t.co/0uQy7nZ5, sekopem & studyreeds
1603	7 desain carroll
1604	yak smp jani https://t.co/GM4nLQp
1605	@NirwinMama lu dia dapat rba 78370 jda lu? ChatGPT?
1606	@gghyhy formidabel chatgpt
1607	RT @AMNabIDGMA: Saya bahkan minta ChatGPT untuk membuatkan saya link lagi tentang Blockchain. LUAR BAYAR. Dari awal sampai akhir dibuat...

Gambar 6. Dataset Hasil Crawling



2. Data Cleaning

Attribute dataset dinormalisasi dengan menghilangkan RT, @, # dan URL, menghasilkan data 754 tweets.



Gambar 7. Operator pada tahap data cleaning

Row No.	Text
1	Kolau di Indonesia stasiun mana yang punya atmosfer komutut
2	Dense dalam 5 tahun lagi kita akan mengalami hal ini
3	Apakah benar kita punya talenta
4	Ya tp cm app ai kayak chatgpt mungkin robot copy ai-dal boleh ka di dm
5	zomaxing...
6	Free je abt Basically dia in browser ChatGPT udah summarize any web page + listing craf app kadi media social
7	PEMANFAATAN ARTIFICIAL INTELLIGENCE UNTUK EFEKTIVITAS PUBLIKASI
8	Rabu 21 Mei 2023
9	Pukul 1:00:00:00 WIB
10	Link register
11	Informasi lebih lanjut hubungi Sekretariat : 0811178239
12	Dosen ITD harap Masyarakat Tak Terlalu Percaya Data Hasil ChatGPT
13	diikuti chatgpt
14	Kali ini bersama team marketing daripada TVET MAWA Sempat tad simple brainstorming guna ChatGPT macan mana rak menarik minat GenZ senazam TVET
15	Inspirasi dalam bengkel nanti akan...
16	Sobaga firdhoter banyak hai yg seger kamu dia kintu ketika pertama kali kerja
17	Jangan takut untuk bertanya dan mencari tahu nnti krn bkrng dan sennntasikan kamu
18	Alau ai knti nnti googling aka bany2 sama ChatGPT

Gambar 8. Dataset setelah cleaning

3. Pelabelan Data

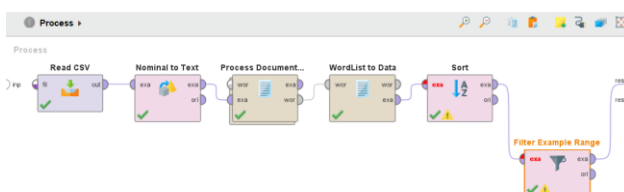
Karena keterbatasan rapidminer terhadap Bahasa Indonesia, maka pelabelan dilakukan mandiri dengan sudut pandang atau penilaian peneliti secara subjektif.

A	B
lgi lagi, lgi lagi, apa yang bikin kejadian berulang ulang kejadian terus? nanya sama chatgpt kali jawabannya ada	positif
Menggunakan Caption Facebook Blisid dengan ChatGPT	positif
Orang kita masih percaya dulun. Kapan majunya ini negara	negatif
Saya cuman bisa senyumin aja ?	positif

Gambar 9. Pelabelan Data

4. Pre-processing Data

Menggunakan dataset yaitu 754 tweets, dengan operator read csv, nominal to text dan melalui text processing. Tahapan tersebut adalah proses persiapan dataset pada text mining.



Gambar 10. Operator pada *pre-processing*

Menggunakan beberapa tahapan proses dengan teknik *TF-IDF* (*Term Frequency — Inverse Document Frequency*) metode algoritma perhitungan bobot setiap kata umum yang digunakan.

a. Tokenize

Fungsi operator ini adalah memisahkan satu kata dengan kata lainnya.

Row No.	A	AB	ABIS	AGAIN	AI	AChatGPT	AKBAR	AKSDHA...	ALLAH	AOSP
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0

Gambar 11. *Tokenize*

WordList (Process Documents from Data)

Word	Attribut...	Total O...	Docum...
A	A	1	1
AB	AB	1	1
ABIS	ABIS	1	1
AGAIN	AGAIN	1	1
AI	AI	68	59
AChatGPT	AChatGPT	1	1
AKBAR	AChatGPT	1	1
AKSDHA...	AKSDHA...	1	1
ALLAH	ALLAH	1	1
AOSP	AOSP	1	1
ARTIFICL...	ARTIFICL...	1	1
ARYSabir...	ARYSabir...	1	1
AS	AS	1	1
ASLI	ASLI	1	1
AU	AU	2	2
Ab	Ab	1	1
Abdul	Abdul	1	1

Gambar 12. *Wordlist Tokenize*

b. *Transform Cases*

Proses penyamarataan font untuk mudah diidentifikasi dalam pemodelan analisis sentimen.

WordList (Process Documents from Data)

Word	Attribut...	Total O...	Docum...
a	a	6	5
ab	ab	2	2
abal	abal	2	1
abul	abul	1	1
abeokuta	abeokuta	1	1
abis	abis	3	3
about	about	2	1
abu	abu	1	1
academic	academic	1	1
academy	academy	1	1
acara	acara	1	1
aceh	aceh	2	1
ada	ada	39	38
adalah	adalah	6	6
adanya	adanya	1	1
add	add	2	2
adjust	adjust	1	1
adminnya	adminnya	1	1

Gambar 13. *Wordlist Transform Cases*



sebanyak 754 tweets yang diolah = 43 (negatif), 711 (positif). Bobot kata yang tidak terdeteksi atau diurutan pertama ada 31, kata 'chatgpt' diurutan kedua yakni 17, dan seterusnya.

Name	Type	Missing	Statistics
data	Polynomial	754	total positif (0) total negatif (0) total positif (0) total negatif (0)
prediction(Sentimen)	Polynomial	0	total positif (43) total negatif (711) total positif (711) total negatif (43)
confidence_positif	Real	0	min 0 max 1 average 0.749
confidence_negatif	Real	0	min 0 max 1 average 0.251
test	Text	0	total zaman ch [:] sdns (1) (31) total (31), chatgpt (17), [562 more]
abal	Real	0	min 0 max 1 average 0.495
abdul	Real	0	min 0 max 1 average 0.345
abeokuta	Real	0	min 0 max 1 average 0.577
abis	Real	0	min 0 max 1 average 0.724

Gambar 19. Hasil Uji Analisis Sentimen Decision Tree

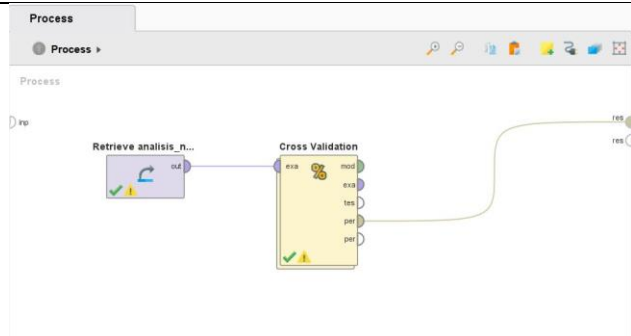
Pada pemodelan analisis sentimen menggunakan naïve bayes, hasil menunjukkan data uji sebanyak 754 tweets yang diolah = 436 (negatif), 318 (positif). Bobot kata yang tidak terdeteksi atau diurutan pertama ada 31, kata 'chatgpt' diurutan kedua yakni 17, dan seterusnya.

Name	Type	Missing	Statistics
data	Polynomial	754	total positif (0) total negatif (0) total positif (0) total negatif (0)
prediction(Sentimen)	Polynomial	0	total positif (318) total negatif (436) total positif (436) total negatif (318)
confidence_positif	Real	0	min 0 max 1 average 0.422
confidence_negatif	Real	0	min 0 max 1 average 0.578
test	Text	0	total zaman ch [:] sdns (1) (31) total (31), chatgpt (17), [562 more]
abal	Real	0	min 0 max 1 average 0.495
abdul	Real	0	min 0 max 1 average 0.345
abeokuta	Real	0	min 0 max 1 average 0.577
abis	Real	0	min 0 max 1 average 0.724

Gambar 20. Hasil Uji Analisis Sentimen Naive Bayes

Perbandingan

Dilakukan perbandingan performa antara algoritma C4.5 dan Naïve bayes, dan dihasilkan sebagai berikut :



Gambar 21. Operator Cross Validation

Didapatkan hasil seperti pada gambar, serta beberapa rincian :

accuracy 77.33%

precision 100.00%

recall 30.18%

Criteria	Table View	Plot View
accuracy	accuracy 77.38% +/- 8.81% (micro average: 77.33%)	
precision		
recall		
AUC (symmetric)		
AUC (asymmetric)		
pred_positif	152	51
pred_negatif	0	22
class_recall	100.00%	30.14%
class_precision	74.88%	

Gambar 22. Hasil Uji Perbandingan

IV. KESIMPULAN

Hasil analisis sentimen telah ditampilkan dalam pengujian, dimana didapat hasil dari pemodelan analisis decision tree sebanyak 754 tweets yang diolah = 43 (negatif), 711 (positif). Bobot kata yang tidak terdeteksi atau diurutan pertama ada 31, kata 'chatgpt' diurutan kedua yakni 17, dan seterusnya. Selanjutnya, hasil dari pemodelan analisis naïve bayes sebanyak 754 tweets yang diolah = 436 (negatif), 318 (positif). Bobot kata yang tidak terdeteksi atau diurutan pertama ada 31, kata 'chatgpt' diurutan kedua yakni 17, dan seterusnya. Perbandingan kedua algoritma mendapat hasil prediksi positif(class precision) = 74.88%, prediksi negatif(class precision) = 100.00%, class recall(true positif) = 100.00%, class recall(true negatif) = 30.14%, yang mana didapat average : accuracy = 77.33%, precision = 100.00%, recall = 30.18%. Kesimpulan dari hasil tersebut adalah positif, artinya respon atau sentimen yang diambil dari sosial media twitter khususnya di Indonesia menunjukkan *feedback* baik terhadap ChatGPT.

V. REFERENSI

- [1] M. Skjuve, A. Følstad, K. I. Fostervold, and P. B. Brandtzaeg, "My Chatbot Companion - a Study of Human-Chatbot Relationships," *International Journal of Human Computer Studies*, vol. 149, May 2021, doi: 10.1016/j.ijhcs.2021.102601.



- [2] N. M. S. Surameery and M. Y. Shakor, "Use Chat GPT to Solve Programming Bugs," *International Journal of Information technology and Computer Engineering*, no. 31, pp. 17–22, Jan. 2023, doi: 10.55529/ijitc.31.17.22.
- [3] Dhiraj. Murthy, *Twitter : social communication in the Twitter age*.
- [4] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 1, p. 149, Jan. 2021, doi: 10.30865/mib.v5i1.2604.
- [5] W. Ertel, "Introduction to Artificial Intelligence," in *Undergraduate Topics in Computer Science*, F. Mast, Ed., 2nd ed. Germany: Springer International Publishing, 2018, pp. 1–93. [Online]. Available: https://www.google.co.id/books/edition/Introduction_to_Artificial_Intelligence/geFHDwAAQBAJ?hl=en&gbpv=1&dq=artificial+intelligence&printsec=frontcover
- [6] S. Skrabut, "80 Ways to Use ChatGPT in the Classroom," in *Computers / Artificial Intelligence / General, Computers / General, Education / General, Artificial intelligence -- Educational applications, Electronic books*, Stan Skrabut, 2023, pp. 1–82. [Online]. Available: https://www.google.co.id/books/edition/80_Ways_to_Use_ChatGPT_in_the_Classroom/ng-rEAAAQBAJ?hl=en&gbpv=1&dq=chatgpt&pg=PT119&printsec=frontcover
- [7] A. Savirani and N. Kurnia, "BIG DATA UNTUK ILMU SOSIAL ANTARA METODE RISET DAN REALITAS SOSIAL," in *Social Science / General, Social Science / Sociology / General*, UGM PRESS, 2021, pp. 11–26. [Online]. Available: https://www.google.co.id/books/edition/BIG_DATA_UNTUK_ILMU_SOSIAL/yHxJEAAAQBAJ?hl=en&gbpv=1&dq=analisi+sentimen&pg=PA213&printsec=frontcover
- [8] A. B. Chaudhuri, "Flowchart and Algorithm Basics," in *Computers / General, Computers / Programming / Algorithms, Computers / Programming / General*, Mercury Learning and Information, 2020, pp. 6–45. [Online]. Available: https://www.google.co.id/books/edition/Flowchart_and_Algorithm_Basics/JJYJEAAAQBAJ?hl=en&gbpv=0
- [9] V. Vyas and V. Uma, "An Extensive study of Sentiment Analysis tools and Binary Classification of tweets using Rapid Miner," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 329–335. doi: 10.1016/j.procs.2017.12.044.
- [10] A. C. Adi, D. P. Lestari, F. S. S. Elsa, and Y. Sabui, "Online School Sentiment Analysis in Indonesia on Twitter Using The Naïve Bayes Classifier and Rapid Miner Tools," *Int. J. Innov. Sci. Res. Technol*, vol. 71, no. 1, pp. 785–788, 2022.
- [11] İstanbul AREL Üniversitesi, IEEE Engineering in Medicine and Biology Society, Institute of Electrical and Electronics Engineers. Turkey Section, and Institute of Electrical and Electronics Engineers, "Sentiment Analysis on University Satisfaction in Social Media".
- [12] W. S. J. Saputra, P. Eva Yulia, and Z. E. Sholikha, "C4.5 and naïve bayes for sentiment analysis Indonesian Tweet on E-Money user during pandemic," in *Proceeding - 6th Information Technology International Seminar, ITIS 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 172–177. doi: 10.1109/ITIS50118.2020.9321081.
- [13] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 765–772. doi: 10.1016/j.procs.2019.11.181.
- [14] F. Albasithu and A. Wibowo, "Perbandingan Algoritma Naïve Bayes Dan C4.5 pada Analisis Sentimen Presiden 3 Periode di Twitter," in *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) Jakarta-Indonesia*, 2022. [Online]. Available: <https://senafiti.budiluhur.ac.id/index.php/>
- [15] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, "Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes," *Information (Switzerland)*, vol. 12, no. 5, May 2021, doi: 10.3390/info12050204.