

Analisis Sentimen Terhadap Aplikasi Chatgpt Pada Twitter Menggunakan Algoritma Naïve Bayes

Muhamad Ilmar Rifaldi¹, Yudhi Raymond Ramadhan.², Irsan Jaelani³

^{1,2,3}Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana Purwakarta, Jawa Barat, Indonesia

e-mail: muhamadilmar08@wastukencana.ac.id¹, yudhi.raymond@wastukencana.ac.id², irsan@wastukencana.ac.id³

Abstract

Sentiment analysis or opinion mining is the detection of attitudes, opinions and emotions towards an object. The value of sentiment analysis can be divided into 2 types of sentiment, namely positive and negative sentiment. One of the topics that is currently being discussed by the public on Twitter social media is an AI-based chatbot application, ChatGPT. This research makes public tweets on Twitter as sentiment analysis material. Data retrieval is done using the Twitter API with data processing on Rapidminer and visualization using Power BI. The sentiment analysis process uses the Naïve Bayes algorithm by dividing the data of 301 tweets into training data and test data. Then testing using Confusion matrix to calculate the performance of the classification results. From the test results, 80% accuracy was obtained, then the Precision value was 80.95% and the Recall value 89.47%. It can be concluded that public sentiment on Twitter for the ChatGPT application tends to be positive seen from the number of positive tweet data as much as 74%.

Keywords: Sentiment Analysis, ChatGPT, Twitter, Confusion Matrix, Naïve Bayes

Abstrak

Analisis sentimen atau opinion mining adalah deteksi terhadap sikap, pendapat maupun emosi terhadap suatu objek. Nilai analisis sentimen dapat dibagi menjadi 2 jenis sentimen yaitu sentimen positif dan negatif. Salah satu topik yang menjadi pembicaraan masyarakat saat ini pada media sosial Twitter adalah sebuah aplikasi chatbot berbasis AI yaitu ChatGPT. Penelitian ini menjadikan tweet masyarakat pada Twitter sebagai bahan analisis sentimen. Penarikan data dilakukan menggunakan Twitter API dengan pengolahan data pada Rapidminer dan visualisasi menggunakan Power BI. Proses analisis sentimen menggunakan algoritma Naïve Bayes dengan membagi data sebanyak 301 tweet menjadi data latih dan data uji. Kemudian pengujian menggunakan Confusion matrix untuk menghitung performance hasil klasifikasi. Dari hasil pengujian didapatkan akurasi sebesar 80%, kemudian nilai Precision sebesar 80,95% dan nilai Recall sebesar 89,47%. Dapat disimpulkan bahwa sentimen masyarakat pada twitter untuk aplikasi ChatGPT cenderung positif dilihat dari jumlah data tweet positif sebanyak 74%.

Kata kunci: Sentimen Analisis, ChatGPT, Twitter, Confusion Matrix, Naïve Bayes

1. PENDAHULUAN

Berkembangnya teknologi informasi berbasis komputer yang pesat berdampak terhadap perubahan dalam berbagai macam pilar kehidupan manusia. kecerdasan buatan (*artificial intelligence*) atau yang lebih dikenal dengan sebutan AI merupakan teknologi terkini produk kemajuan teknologi yang pesat. AI membuat komputer dapat melakukan banyak pekerjaan seperti yang dilakukan oleh manusia sehingga menjadi produk teknologi yang banyak dimanfaatkan dalam pembuatan aplikasi pada saat ini karena memudahkan manusia dalam memenuhi kebutuhannya[1].

Salah satu produk teknologi berbasis *Chat* dengan dukungan AI adalah *Chatbot*, yaitu merupakan aplikasi komputer yang dibuat untuk berinteraksi dengan orang melalui pesan teks dan suara. Mereka memiliki kecerdasan buatan dan pemrosesan bahasa alami (NLP) untuk menjadi aplikasi komputer yang cerdas yang dapat menjawab pertanyaan manusia.[1] ChatGPT (*Generative Pre-trained Transformer*) dari OpenAI merupakan teknologi informasi dan komunikasi berbasis kecerdasan buatan yang telah dilatih untuk meniru percakapan manusia melalui teknologi pemrosesan bahasa alami atau NLP. Pada pengaplikasiannya ChatGPT dapat menghasilkan suatu tulisan yang bersifat ilmiah atau bahkan baku dengan teknik yang baik dan efektif dengan *prompt* yang dirumuskan di awal [2].

Zhai, seorang ahli kecerdasan buatan, melakukan eksperimen dengan menggunakan ChatGPT untuk membuat artikel berjudul "*Artificial Intelligence for Education*" pada tahun 2023. Dia menilai artikel buatan mesin itu bersifat koheren, relatif akurat, informatif, dan sistematis. Chatbot dinilai lebih baik dan efisien dalam memberikan informasi daripada manusia biasa dan lebih baik dalam menulis sebuah karya ilmiah daripada rata-rata mahasiswa pada umumnya.[2].

Dunia pendidikan juga merespons ketika melihat kemampuan ChatGPT untuk membuat tulisan yang terorganisir dengan baik. Pada 12 Desember 2022, *Los Angeles Unified School District* menghentikan akses ke website OpenAI ChatGPT pada jaringan dan perangkat lunak sekolah di distrik mereka. Pada akhir bulan Desember 2022, Departemen Pendidikan Kota New York juga melakukan hal yang sama terhadap sekolah-sekolah di distrik tersebut. Larangan ChatGPT diberikan karena itu tidak membantu siswa belajar berpikir kritis (*critical thinking*) dan memecahkan masalah (*problem solving*). Kedua kemampuan ini penting untuk kesuksesan mereka di sekolah dan untuk bekal masa depan.[2].

Kemampuan luar biasa yang dimiliki aplikasi *ChatGPT* ini menimbulkan polemik pada masyarakat dikalangan akademis, karena dianggap bisa menimbulkan ketergantungan penggunaan yang berlebihan juga dapat menurunkan gairah penulisan karya ilmiah karena kemudahan yang ditawarkan oleh aplikasi ChatGPT ini dalam membuat artikel selevel jurnal dengan baik. Oleh karena itu untuk mengetahui sentimen masyarakat terhadap teknologi ChatGPT ini diperlukan sebuah analisis yang dapat mengklasifikasikan sentimen negatif maupun positif yang ada pada masyarakat. Analisis sentimen adalah proses menganalisis pendapat atau opini, evaluasi, sikap, penilaian, dan emosi seseorang, terlepas dari apakah pembicara atau penulis berbicara tentang subjek, produk, layanan, organisasi, individu, atau kegiatan tertentu.[3]. Persiapan pencarian model yang menggambarkan data dan mengklasifikasikannya pada kelas-kelas yang tersedia dikenal sebagai teknik klasifikasi. Karena unggul dalam kecepatan dan akurasi, Naive Bayes sebagai salah satu algoritma klasifikasi data dianggap sebagai metode klasifikasi data terbaik dalam hal akurasi dan perhitungan.[4].

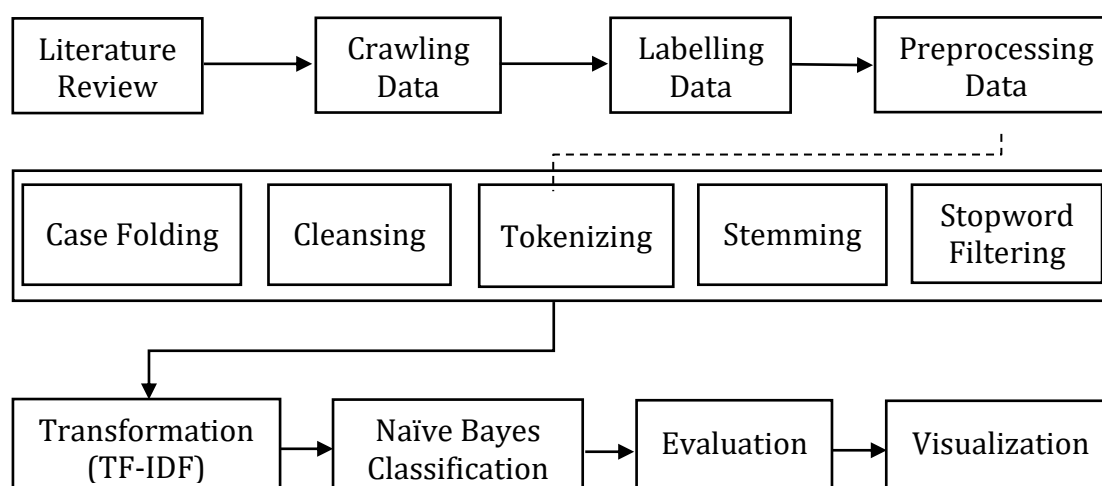
Twitter adalah salah satu *platform* media sosial yang banyak dipakai oleh masyarakat Indonesia. Karena kepopulerannya, menjadikan Twitter salah satu laman untuk mengekspresikan sentimen publik mengenai isu yang sedang dalam perbincangan. Pada media sosial ini, masyarakat membicarakan berbagai macam

hal yang sedang terjadi, baik di bidang teknologi, pendidikan, ekonomi, politik, sosial budaya, hingga isu-isu viral ditengah masyarakat tidak terkecuali mengenai aplikasi ChatGPT[5].

Berdasarkan pemaparan diatas maka penelitian ini bertujuan untuk mengetahui sentimen masyarakat terhadap aplikasi chatgpt pada media sosial twitter berupa sentimen positif dan negatif dengan menggunakan data *tweet* yang beredar dengan metode klasifikasi menggunakan algoritma Naïve Bayes hingga kemudian dapat dinilai apakah sentimen masyarakat cenderung kearah positif atau negatif.

2. METODOLOGI PENELITIAN

Penelitian ini dilakukan dengan metode klasifikasi menggunakan algoritma Naïve Bayes. langkah-langkah yang dilakukan yaitu *Literature Review*, *Crawling Data*, *Labelling Data*, *Preprocessing Data*, *Transformation*, *Naïve Bayes Classification*, *Evaluation* dan *Visualization*. Untuk kerangka penelitian dapat dilihat pada Gambar 1 dibawah ini:



Gambar 1. Metodologi Penelitian

Penjelasan untuk tahapan pada penelitian ini adalah sebagai berikut:

- 1) *Literature Review* yaitu proses untuk melakukan pengumpulan data sebagai referensi berupa jurnal, artikel, atau karya ilmiah lainnya yang dapat menunjang penelitian ini mengenai data mining, sentimen analisis, klasifikasi naïve bayes, dan ChatGPT.
- 2) *Crawling Data* merupakan tahapan yang bertujuan untuk mengumpulkan atau mengambil data dari sebuah database.[6] pada tahap ini data diunduh pada platform twitter dengan menggunakan API berupa *tweet* yang berhubungan dengan ChatGPT. Proses ini dilakukan menggunakan *tools* Rapidminer.

- 3) *Labelling Data* adalah tahap dimana data yang telah dikumpulkan akan diberikan label. Kategori yang digunakan dalam penelitian ini terdiri atas label positif bagi data yang mengandung sentimen positif dan label negatif untuk data yang mengandung sentimen negatif.[7]
- 4) *Preprocessing* yaitu proses pembersihan dan penyiapan teks untuk analisis data dari berbagai komponen yang tidak informatif seperti *URL*, tanda baca, dan karakter lainnya.[8] pada tahap ini data akan dibersihkan dari atribut-atribut yang tidak diperlukan sehingga menghasilkan data yang siap untuk diproses untuk analisis sentimen.
- 5) *Transformation* pada tahap ini untuk mendapatkan nilai dari setiap kata yang ada data di transformasikan dengan memperhitungkan bobot kata pada tiap dokumen *tweet* sehingga dapat diketahui kemiripan suatu kata pada dokumen *tweet* lainnya. pada penelitian ini pembobotan kata dilakukan yakni dengan metode TF-IDF yang merupakan suatu metode yang menentukan seberapa jauh keterkaitan kata terhadap dokumen dengan memberikan bobot atau nilai pada setiap kata.[9]
- 6) *Naïve Bayes Classification* merupakan pengklasifikasian dengan metode statistika yang dapat digunakan untuk memprediksi peluang keanggotaan kelas, seperti probabilitas tupel tertentu untuk bergabung pada kelas tertentu. Metode ini mengasumsikan semua atribut independen dan dalam pengaplikasiannya memberikan tingkat akurasi dan kecepatan yang cukup tinggi. Naïve Bayes terbukti efektif dalam pengklasifikasian *text mining*. [10] Pada proses pengujian, data dibagi secara acak menjadi dua kategori yaitu data pelatihan dan data pengujian. Data pelatihan adalah data yang dikumpulkan dan digunakan untuk membuat model yang dapat memprediksi apakah *tweet* termasuk dalam kategori positif atau negatif. Sedangkan data pengujian adalah data yang digunakan untuk menguji akurasi model klasifikasi. Pada penelitian ini, rasio antara data uji dan data latih adalah 80 banding 20.
- 7) *Evaluation* pada tahap ini hasil klasifikasi dievaluasi menggunakan *Confussion Matrix*, dimana ia dapat menampilkan jumlah prediksi benar dan salah yang dihasilkan oleh model dibandingkan dengan hasil klasifikasi sebenarnya pada data uji.[11] pengukuran dengan *confussion matrix* meliputi *Accuracy*, *Precision*, dan *Recall*.
- 8) *Visualization* pada tahap ini hasil penelitian divisualisasikan agar mempermudah untuk memahami hasilnya, pada penelitian ini visualisasi disajikan dalam bentuk *Pie chart* menggunakan *tools* Power BI.

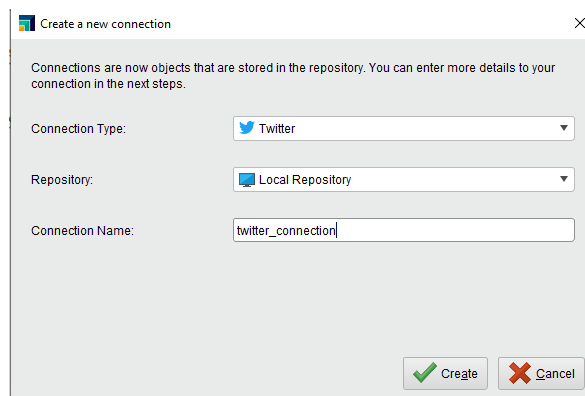
3. HASIL DAN PEMBAHASAN

Pada tahap ini akan dipaparkan hasil dari analisis dengan menggunakan metodologi yang telah dijelaskan sebelumnya meliputi *Crawling Data*, *Labelling Data*, *Preprocessing Data*, *Transformation*, *Naïve Bayes Classification*, *Evaluation*, dan *Visualization*.

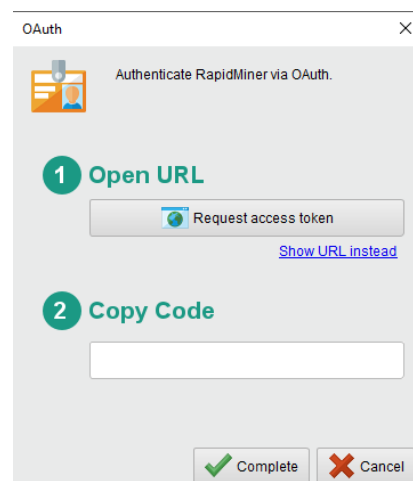
3.1. *Crawling Data*

Data diambil menggunakan *API (Application Programming Interface)* Twitter dengan bantuan *tools* Rapidminer pada tanggal 20 Mei 2023. Proses pengambilan data diawali dengan membuat koneksi terhadap Twitter dengan menggunakan *API* pada *tools* Rapidminer. Proses *create connection* dapat dilihat pada Gambar 2. Selanjutnya dengan *tools* Rapidminer kita dapat dengan langsung *me-request* autentikasi untuk mendapat akses token kepada Twitter dengan terlebih dulu sudah login menggunakan akun Twitter yang ada. Proses *request* autentikasi akses token dapat dilihat pada Gambar 3.

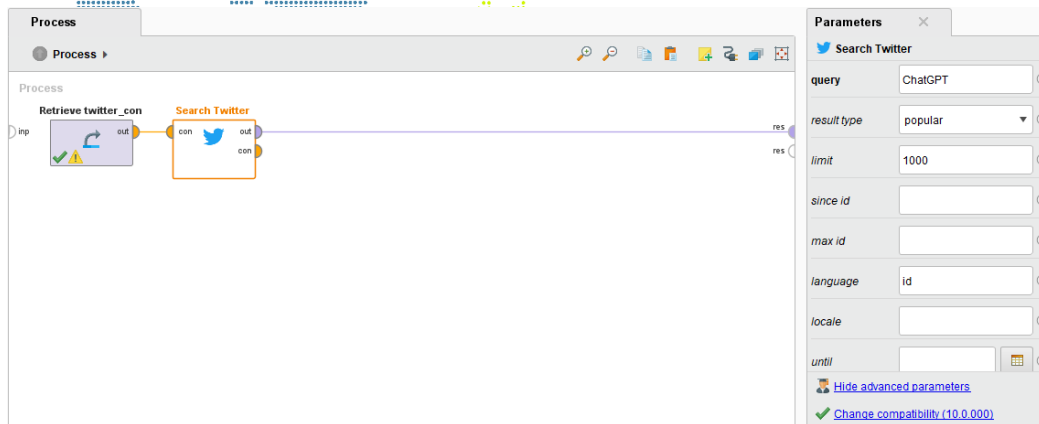
Lalu dengan menggunakan operator *Search Twitter* dapat ditentukan *query* pencarian yang diinginkan dan jumlah tweet yang akan di *crawling*. Hasil dari *crawling data* mendapatkan 1000 *tweet* dengan *query* yang digunakan yaitu ChatGPT. Proses *Input query* pencarian *tweet* dapat dilihat pada Gambar 4.



Gambar 2. Membuat koneksi Twitter



Gambar 3. Request akses token



Gambar 4. Input Query Pencarian tweet

Setelah proses Crawling selesai hasil yang diperoleh dapat dilihat pada Gambar 5 dibawah ini.

Created At Date / Time	From-User Category	From-User-Id Category	To-User Category	To-User-Id Category	Language Category	Source Category	Text Category	Geo-Location Number	Geo-Lo Number
May 18, 2023 3:...	Ainun Najib	35167068	?	-1	in	<a href="https://...	Bard beneran ...	?	?
May 19, 2023 1...	Pelabur Bijak	118794822824...	?	-1	in	<a href="http://t...	Guna VR + Cha...	?	?
May 21, 2023 1...	Irham.	790924465932...	bleuivy	758775811239...	in	<a href="http://t...	@bleuivy bener...	?	?
May 21, 2023 1...	ta'	150016944118...	?	-1	in	<a href="http://t...	aku kemarin tu ...	?	?
May 21, 2023 1...	MIRAH	127204486250...	?	-1	in	<a href="http://t...	RT @myskill_j...	?	?
May 21, 2023 1...	pua \ having a ...	134537105946...	Bluemberryz	163188420588...	in	<a href="http://t...	@Bluemberryz...	?	?
May 21, 2023 1...	acil, dm yg di tag	894051541228...	bigwinjuseyyong	143512812794...	in	<a href="http://t...	@bigwinjusejy...	?	?
May 21, 2023 1...	Irham.	790924465932...	zorzaith	2249420755	in	<a href="http://t...	@zorzaith usa...	?	?
May 21, 2023 1...	Kuan wangy	126159290856...	?	-1	in	<a href="http://t...	RT @myskill_j...	?	?

Gambar 5. Hasil Crawling Data

3.2. Labelling Data

Pada tahap ini data yang telah didapat akan diberikan label secara manual dengan menggunakan Microsoft excel yaitu dengan label positif dan negatif. Pelabelan dilakukan dengan manual agar hasil pelabelan sentimen dapat menjadi lebih akurat dengan penilaian langsung dari manusia. Hasil pelabelan manual sentimen dapat dilihat pada Tabel 1 dibawah ini:

Tabel 1. Hasil pelabelan manual

Sentimen	Text	user
Positif	Terimakasih chatgpt sangat menolong sekali:) https://t.co/GMd9nJZ2cr	azeggg
Negatif	chatGPT makin meresahkan	_Xulinho

Setelah proses pelabelan data, didapatkan sebanyak 301 data yang mengandung sentimen negatif dan positif. Jumlah data *tweet* positif dan negatif dapat dilihat pada Tabel 2 dibawah ini:

Tabel 2. Jumlah tweet positif dan negatif

Tweet	Jumlah Data
Positif	223
Negatif	78

3.3. Preprocessing Data

Selanjutnya tahap *preprocessing* yaitu dimana data diproses sedemikian rupa sehingga data akan siap untuk diolah. Adapun hasil yang telah didapat dalam prosesnya yaitu:

3.3.1. Case Folding

Semua teks atau kalimat dalam data *tweet* diubah menjadi huruf kecil (*lower case*) melalui proses *case folding*, yang memudahkan untuk mencocokkan dokumen dengan ukuran huruf yang sama.[4] hasil dari proses ini dapat dilihat pada Tabel 3 dibawah ini:

Tabel 3. Hasil proses *Case Folding*

Sebelum
AKU DAH GABISA HIDUP RASANYA TANPA CHATGPT SEKARANG
Sesudah
aku dah gabisa hidup rasanya tanpa chatgpt sekarang

3.3.2. Cleansing

Proses cleansing dilakukan untuk membersihkan data dari karakter-karakter yang tidak dibutuhkan seperti angka, tanda baca, *URL*, emotikon, dan lainnya.[12] hasil dari proses ini dapat dilihat pada Tabel 4 dibawah ini:

Tabel 4. Hasil proses *Cleansing*

Sebelum
ada chatgpt aku malah makin manja.. bales email aja minta tolong dia .-. tapi emang pinter banget sih, cinta!! https://t.co/99b16DGjkb
Sesudah
ada chatgpt aku malah makin manja bales email aja minta tolong dia tapi emang pinter banget sih cinta

3.3.3. Tokenizing

Proses ini melakukan pemecahan suatu teks berdasarkan setiap penyusun katanya, sehingga menjadi bagian-bagian kecil yang disebut token.[5] hasil dari proses *tokenizing* dapat dilihat pada Tabel 5 dibawah ini:

Tabel 5. Hasil proses *Tokenizing*

Sebelum				
ada juga yang tidak bisa dijawab oleh chatgpt wkwkwk				
Sesudah				
Ada	juga	yang	tidak	bisa
dijawab	oleh	chatpt	wkwkwk	

3.3.4. Stemming

Merupakan tahap dimana kata yang memiliki imbuhan dihilangkan hingga kata tersebut menjadi kata dasarnya.[13] Pada tahap ini, kata-kata yang berimbuhan akan diganti menjadi kata dasar, Hasil dari proses *stemming* dapat dilihat pada Tabel 6 dibawah ini:

Tabel 6. Hasil proses *Stemming*

Sebelum
chatgpt cukup membantu banget dalam menyusun kata
Sesudah
chatgpt cukup bantu banget dalam susun kata

3.3.5. Stopwords Filtering

Pada tahap ini kata yang dianggap tidak penting akan dihapus agar dapat meningkatkan kecepatan pemrosesan data, hanya kata relevan dan memiliki makna penting yang diambil.[14]. hasil dari proses *Stopwords Filtering* dapat dilihat pada Tabel 7 dibawah ini:

Tabel 7. Hasil proses *Stopwords Filtering*

Sebelum
ah relate banget pakai chatgpt aja hehe aku begitu kok hampir semua malah cuma beberapa bagian aja yang aku paraphrase sendiri biar jadi lebih punya emosi gitu
Sesudah
relate pakai chatgpt beberapa bagian yang paraphrase sendiri lebih punya emosi

3.4. Transformation (TD-IDF)

Tahap berikutnya yaitu proses Transformasi data menggunakan metode TF-IDF Dimana setiap kata akan dihitung bobotnya. *Term Frequency-Inverse Document Frequency (TF-IDF)* merupakan suatu metode transformasi yang menentukan keterkaitan kata terhadap dokumen dengan memberikan bobot pada setiap kata.[9] Rumus dari TF-IDF dapat dilihat dibawah ini:

$$IDF (Word) = \text{Log} \frac{td}{df} \tag{1}$$

$$TF-IDF = TF \times IDF \tag{2}$$

Perhitungan *Term Frequency* dilakukan pada 10 sampel dokumen dengan 8 kata pilihan yaitu chatgpt, solusi, membantu, jelek, mudah, *error*, curhat, dan ngelag. Term yang muncul diberi nilai 1 dan jika tidak muncul diberi nilai 0. Berikut hasil perhitungan *term frequency* dapat dilihat pada Tabel 8.

Tabel 8. Hasil perhitungan *Term Frequency*

Term	TF									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10

Term	TF									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
chatgpt	1	1	1	1	1	1	1	1	1	1
solusi	0	1	1	0	0	0	0	0	0	0
membantu	0	0	0	0	1	0	0	0	0	0
jelek	0	0	0	0	0	0	0	0	0	1
mudah	1	0	0	0	0	0	0	0	0	0
eror	0	0	0	0	0	0	1	0	0	0
curhat	0	0	1	0	0	0	0	0	0	0
ngelag	0	0	0	0	0	1	0	0	0	0

Kemudian menghitung jumlah frekuensi kata terpilih yang sudah diketahui, kemudian menghitung jumlah dokumen (IDF), di mana IDF adalah hasil inverse dari DF. Hasil IDF disajikan pada Tabel 9.

Tabel 9. Hasil perhitungan IDF

DF	D/DF	LOG(D/DF)	1+LOG
10	1	0	1
2	5	0,69897	1,69897
1	10	1	2
1	10	1	2
1	10	1	2
1	10	1	2
1	10	1	2
1	10	1	2

Berikutnya dapat dihitung hasil TF-IDF dengan mengalikan hasil TF dan IDF, hasilnya dapat dilihat pada Tabel 10.

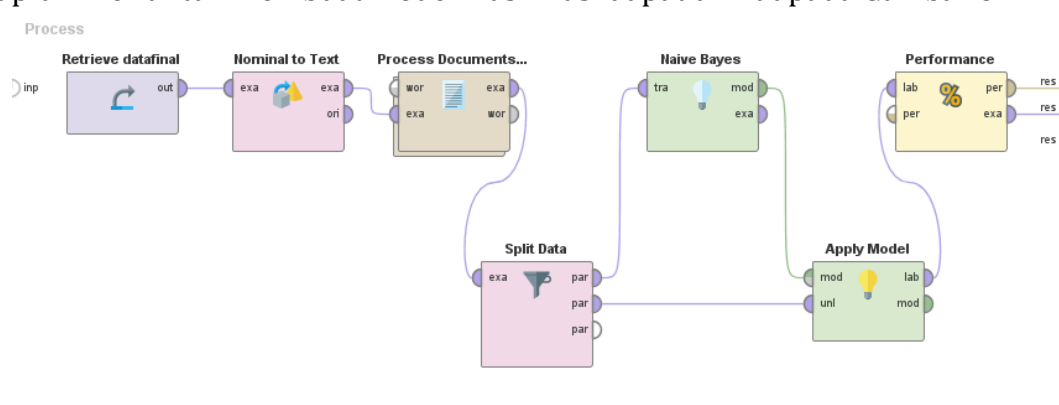
Tabel 10. Hasil perhitungan TF-IDF

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	1	1	1	1	1	1	1	1	1
0	1,69897	1,69897	0	0	0	0	0	0	0
0	0	0	0	2	0	0	0	0	0
0	0	0	0	0	0	0	0	0	2
2	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	2	0	0	0
0	0	2	0	0	0	0	0	0	0

3.5. Naïve Bayes Classification

Pengklasifikasian menggunakan metode Naïve Bayes menggunakan Rapidminer diawali dengan mengimport data pada Rapidminer. Lalu data akan

dirubah dari nominal kedalam *text*. Selanjutnya data akan melalui proses *preprocessing text*, sehingga menghasilkan data yang telah ditransformasi dengan metode TF-IDF. Lalu data yang telah ditransformasi akan dibagi menjadi data latih dan data uji, setelah itu diklasifikasikan dengan menggunakan algoritma Naïve Bayes. Kemudian algoritma Naïve Bayes akan diaplikasikan pada data dan selanjutnya akan dihitung performa modelnya. Rancangan pemodelan pada Rapidminer untuk membuat model klasifikasi dapat dilihat pada Gambar 6.



Gambar 6. Model klasifikasi Naive Bayes

Setelah dilakukan klasifikasi pada data menggunakan Rapidminer dengan algoritma Naive Bayes kinerja model akan menghasilkan nilai performa. Berikut ini adalah hasil *performance* model dapat dilihat pada Gambar 7.

Criterion: accuracy

Table View Plot View

accuracy: 80.00%

	true POSITIF	true NEGATIF	class precision
pred. POSITIF	34	8	80.95%
pred. NEGATIF	4	14	77.78%
class recall	89.47%	63.64%	

Gambar 7. Hasil Klasifikasi Naive Bayes

Kinerja dari klasifikasi sentimen menggunakan algoritma Naïve Bayes menggunakan metode *Split Validation* dengan rasio 80:20 dimana 80% adalah sebagai data latih dan 20% sebagai data uji secara keseluruhan menghasilkan tingkat akurasi sebesar 80,00%, dengan nilai *Precision* untuk kategori positif sebesar 80,95% dan untuk kategori negatif sebesar 77,78%. Lalu nilai *Recall* untuk kategori Positif sebesar 89,47% dan kategori negatif sebesar 63,64%.

3.6. Evaluation

Setelah proses klasifikasi dengan Naïve Bayes dilakukan, tahap berikutnya yaitu proses evaluasi yang bertujuan untuk mengetahui nilai akurasi, *precision*,

dan *recall* dari hasil yang didapatkan. Proses ini menggunakan Tabel *Confussion Matrix*. Berikut adalah *Confussion matrix* dapat dilihat pada Tabel 11.

Tabel 11. Hasil *Confussion matrix*

Prediksi	Aktual	
	Positif	Negatif
Positif	34	8
Negatif	4	14

Dari Tabel 11 dapat dihitung secara manual untuk nilai *Accuracy*, *Precision*, dan *Recall*. Perhitungan dapat dilihat dibawah ini:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{34+14}{34+14+8+4} = 0,8 \times 100\% = 80\% \quad (3)$$

$$Precision = \frac{TP}{TP+FP} = \frac{34}{34+8} = 0,8095 \times 100\% = 80,95\% \quad (4)$$

$$Recall = \frac{TP}{TP+FN} = \frac{34}{34+4} = 0,8947 \times 100\% = 89,47\% \quad (5)$$

Keterangan:

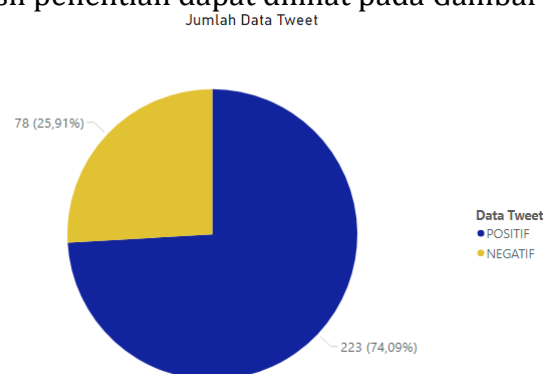
Accuracy : Penilaian kinerja klasifikasi secara keseluruhan

Precision : Prediksi benar positif dari hasil prediksi positif

Recall : Prediksi benar positif dari populasi

3.7. Visualization

Untuk mempermudah memahami hasil pembahasan, maka diperlukan visualisasi data. Visualisasi data menggunakan *Pie Chart* untuk melihat jumlah keseluruhan *tweet* yang berisi sentimen positif dan negatif. Berikut ini hasil visualisasi data untuk hasil penelitian dapat dilihat pada Gambar 8.



Gambar 8. Pie Chart jumlah *tweet* positif dan negatif

4. SIMPULAN

Dari hasil penelitian dengan menggunakan data sebanyak 301 *tweet* pengklasifikasian data dengan metode Naïve Bayes dan pengujian menggunakan *Confussion Matrix* didapatkan hasil akurasi sebesar 80%. Dengan nilai *Precision*

sebesar 80,95% dan nilai *Recall* sebesar 89,47%. Dapat dinilai bahwa analisis sentimen dengan algoritma Naïve bayes mampu mengklasifikasikan sentimen pada twitter dengan baik dengan memperoleh akurasi sebesar 80%. Juga dapat disimpulkan bahwa sentimen masyarakat terhadap aplikasi ChatGPT pada media sosial Twitter cenderung positif dengan jumlah presentasi *tweet* sebanyak 74,09%. Hasil penelitian ini dapat menjadi acuan bagi pemerintah khususnya bagi kementerian informasi dan komunikasi yang bertanggung jawab untuk memfilter aplikasi yang beredar pada masyarakat, bahwasanya aplikasi ChatGPT ini cenderung diterima oleh masyarakat Indonesia melihat lebih banyaknya sentimen positif dari masyarakat pada Twitter. Sehingga tidak perlu dilakukan pelarangan atau pemblokiran terhadap ChatGPT. Sebagai saran untuk penelitian selanjutnya pada kasus ini dapat menggunakan algoritma lainnya seperti *Support Vector Machine*, K-NN, dan DBSCAN. Lalu dapat diaplikasikan dengan algoritma optimalisasi seperti *Particle Swarm Optimization* sehingga dapat menghasilkan referensi karya ilmiah lainnya sebagai bahan pengembangan keilmuan khususnya dalam bidang *Data Science*.

DAFTAR PUSTAKA

- [1] Mukrodin and N. Mega Sasmita, "rtificial Inteligence Dalam Apilkasi Chatbot Sebagai Helpdesk Obyek Wisata Dengan Permodelan Natural Language Processing (Studi Kasus: Kabupaten Cilacap)," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 10, no. 1, pp. 7–14, 2021, doi: 10.30591/smartcomp.v10i1.2135.
- [2] A. Setiawan and U. K. Luthfiyani, "Penggunaan ChatGPT Untuk Pendidikan di Era Education 4.0: Usulan Inovasi Meningkatkan Keterampilan Menulis," *J. PETISI*, vol. 04, no. 01, 2023, [Online]. Available: <https://chat.openai.com>.
- [3] N. Sucahyo, I. Kuniati, and K. Harvit, "Analisis Sentimen Masyarakat Terhadap UU Cipta Kerja Pada Media Sosial Twitter," vol. 02, no. 01, 2022.
- [4] E. Undamayanti *et al.*, "Analisis Sentimen Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 2, pp. 916–930, 2022.
- [5] T. N. Wijaya, R. Indriati, and M. N. Muzaki, "Analisis Sentimen Opini Publik Tentang Undang-Undang Cipta Kerja Pada Twitter," *Jambura J. Electr. Electron. Eng.*, vol. 3, no. 2, pp. 78–83, 2021, doi: 10.37905/jjee.v3i2.10885.
- [6] A. Anugrah, T. I. Hermanto, and I. Kaniawulan, "Sentiment Analysis Of Internet Service Providers Using Naïve Bayes Based On Particle Swarm Optimization," *J. Ris. Inform.*, vol. 4, no. 4, pp. 371–378, 2022, doi: 10.34288/jri.v4i4.408.
- [7] A. Nabillah, S. Alam, and M. G. Resmi, "Twitter User Sentiment Analysis Of TIX ID Applications Using Support Vector Machine Algorithm," vol. 3, no. 1, pp. 14–27, 2022.
- [8] A. E. Augustia, R. Taufan, Y. Alkhalifi, and W. Gata, "Analisis Sentimen Omnibus Law Pada Twitter Dengan Algoritma Klasifikasi Berbasis Particle Swarm Optimization," *Paradig. - J. Komput. dan Inform.*, vol. 23, no. 2, pp. 158–166, 2021, doi: 10.31294/p.v23i2.10430.
- [9] D. A. Wulandari and R. Saedudin, Rd. Rohmat Andreswari, "Analisis Sentimen Media Sosial Twitter Terhadap Reaksi Masyarakat Pada Ruu Cipta Kerja Menggunakan

- Metode Klasifikasi Algoritma Naive Bayes Analysis,” vol. 8, no. 5, pp. 9007–9016, 2021.
- [10] N. Helmiah *et al.*, “Penerapan Metode Naive Bayes dalam Analisis Persepsi Masyarakat mengenai Rencana Pengesahan RUU Omnibus Law di Bidang Investasi dan Ketenagakerjaan Tahun 2020 di Indonesia,” *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 8, no. 2, p. 48, 2020, doi: 10.24252/msa.v8i2.16743.
- [11] I. K. Syahputra, F. A. Bachtiar, and S. A. Wicaksono, “Implementasi Data Mining untuk Prediksi Mahasiswa Pengambil Mata Kuliah dengan Algoritme Naive Bayes,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 5902–5910, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3464>
- [12] I. G. N. Daffa Adnyana, F. Adams, and A. Windari Oktavia, “Analisis Sentimen Terhadap Undang-Undang Cipta Kerja Menggunakan Metode Naive Bayes,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl. Jakarta-Indonesia*, no. September, pp. 120–129, 2021.
- [13] D. Siti Utami and A. Erfina, “Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, pp. 418–427, 2022.
- [14] Y. Nurdiansyah, F. Rahman, and P. Pandunata, “Analisis Sentimen Opini Publik Terhadap Undang-Undang Cipta Kerja pada Twitter Menggunakan Metode Naive Bayes Classifier,” *Pros. Semin. Nas. Sains Teknol. dan Inov. Indones.*, vol. 3, no. November, pp. 201–212, 2021, doi: 10.54706/senastindo.v3.2021.158.