

Analisis Sentimen Persepsi Masyarakat Terhadap Timnas Indonesia U-23 dalam AFC-23 Asian Cup 2024 Pada Media Sosial X Menggunakan Metode *Naïve Bayes Classifier*

Reda Hayati Nufus¹, Untung Surapati^{2*}

^{1,2*} Program Studi Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

Email: nufusredha@gmail.com¹, kisuro2003@gmail.com^{2*}

Histori Artikel:

Dikirim 26 Mei 2024; *Diterima dalam bentuk revisi* 25 Juni 2024; *Diterima* 17 Juli 2024; *Diterbitkan* 20 September 2024. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Media sosial Twitter atau sekarang disebut dengan X tidak hanya berfungsi sebagai media komunikatif yang efektif, namun juga dapat memudahkan masyarakat mencari informasi atau berinteraksi langsung dengan pengguna X lainnya, dan menjadi wadah untuk menampung opini-opini. Pengguna X di Indonesia sering mengutarakan pendapatnya pada semua bidang termasuk olahraga, khususnya Timnas Sepakbola Indonesia yang baru-baru ini sedang hangat diperbincangkan karena bakat semua pemain yang berhasil menarik perhatian tidak hanya penggemar bola saja, namun seluruh masyarakat Indonesia pada kompetisi AFC Asian Cup U-23 2024 yang dilaksanakan di negara Qatar. Penelitian ini menggunakan metode klasifikasi Naïve Bayes untuk menganalisa sentimen komentar-komentar pada Sosial Media X terhadap setiap pertandingan. Data sentimen dari komentar pengguna pada sosial media X dikumpulkan dan diklasifikasikan sebagai positif, dan negatif. Hasil akhir dari analisis sentimen dengan metode naïve bayes menunjukkan nilai akurasi sebesar 73,94% dari 972 dataset, terdapat 143 data terprediksi negatif dan 828 data yang terprediksi Positif. Sehingga dapat disimpulkan bahwa sebagian besar persepsi masyarakat positif karena melihat perjuangan dan kinerja Timnas Sepakbola Indonesia dalam mengharumkan nama Bangsa.

Kata Kunci: Media Sosial X; Timnas Sepakbola Indonesia; Algoritma Naïve Bayes.

Abstract

Social media Twitter, or now called X, not only serves as an effective communicative medium, but it can also make it easier for people to find information or interact directly with other X users, and become a container for holding opinions. X users in Indonesia often express their opinions on all areas including sports, especially the Indonesian Football Team, which has recently been heatedly discussed because of the talent of all the players who succeeded in attracting the attention not only of football fans, but of the entire Indonesia public at the AFC ASIA CUP U 23 2024 competition held in the country of Qatar. The study uses the Naïve Baiyes classification method to analyze the sentiment of comments and comments on Social Media X against each match. Sentiment data from user comments on social media X is collected and classified as positive, and negative. The final results of sentimental analysis using the naïve bayes method showed accuracy values of 73.94% out of 972 datasets, with 143 predicted negative data and 828 predictable positive data. So it can be concluded that most of the public's perception is positive because they see the struggle and performance of the Indonesian National Football Team in making the nation proud.

Keyword: Social Media X; The Indonesian Football Team; Algorithm Naïve Bayes.

1. Pendahuluan

Pada era digital ini, media sosial telah menjadi salah satu platform utama bagi masyarakat untuk berbagi pendapat, opini, dan sentimen mereka terhadap berbagai topik, termasuk olahraga, salah satunya *Twitter* atau sekarang dikenal dengan sebutan *X*. *X* merupakan salah satu aplikasi media sosial berbasis *website* dan *mobile* yang sering digunakan sebagai alat komunikasi (Sumitro, Mulyana, & Saputro, 2021). Salah satu fitur *X* yang paling sering digunakan adalah fitur *Tweet*, yang memungkinkan pengguna untuk menuliskan pemikiran, pandangan, dan pendapatnya (Akbar & Ihsan, 2023). Pada saat yang sama, analisis sentimen telah menjadi alat yang efektif dalam memahami dan mengukur respons masyarakat terhadap suatu topik atau entitas tertentu (Matheos Sarimole & Kudrat, 2024). Hal yang sedang menjadi topik utama dan banyak dikomentari oleh pengguna *X* di Indonesia baru-baru ini adalah mengenai Timnas Indonesia U-23 dalam partisipasinya di kompetisi AFC U-23 Asian Cup 2024 yang dilaksanakan di negara Qatar. Telah ramai menjadi perhatian utama bagi banyak individu dan komunitas di media sosial, terlebih lagi Timnas Indonesia memiliki penggemar yang kuat di Indonesia (Sulastiyono, Setiawan, & Nugroho, 2023). Analisis terhadap sentimen terkait Timnas Indonesia U-23 tidak hanya dapat mengungkapkan popularitas mereka, tetapi juga bagaimana masyarakat memandang kinerja dan hasil mereka di kompetisi internasional. Metode yang digunakan adalah *Naïve Bayes Classifier* yang telah terbukti efektif dalam mengklasifikasikan teks berdasarkan sentimen yang terkandung di dalamnya.

Sebelumnya, sudah banyak penelitian yang menggunakan metode *Naïve Bayes* untuk melakukan analisis sentimen. Seperti penelitian yang dilakukan oleh Ardianne Luthfika Fairuz, Rima Dias Ramadhani, dan Nia Annisa Ferani Tanjung (2021) pada Jurnal DINDA, yang menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* untuk membangun model *sentiment analysis* terhadap *tweet* pengguna *Twitter* terkait COVID-19. Kedua algoritma ini bekerja dengan baik, dengan hasil akurasi mencapai lebih dari 72%. Hasil akurasi yang didapatkan oleh algoritma *Naïve Bayes* lebih baik dibandingkan *K-Nearest Neighbor*, yaitu 85%, sementara pada algoritma *K-Nearest Neighbor* didapatkan angka akurasi sebesar 82%. Penelitian lain yang dilakukan oleh Rini Maria, Retno Umi Umayah, Syifa Mahardanny, Diki Nur Kalana, dan Dedi Dwi Saputra (2023) pada Jurnal Komputer Antartika, juga menggunakan metode *Naïve Bayes Classifier* untuk menganalisis sentimen masyarakat terhadap penggunaan Aplikasi *My Pertamina* di media sosial *Twitter*. Penelitian ini bertujuan untuk mengetahui kecenderungan komentar atau *tweets* pengguna *Twitter* terhadap kebijakan pembelian Peralite menggunakan Aplikasi *MyPertamina*. Hasil penelitian menunjukkan bahwa nilai *accuracy* sebesar 84,08%, nilai *precision* sebesar 84,51%, nilai *recall* sebesar 83,81%, dan nilai *AUC* sebesar 0,903 yang dievaluasi dengan *confusion matrix*.

Dengan menggunakan pendekatan tersebut, penelitian ini mampu mengidentifikasi pola sentimen yang muncul di media sosial terhadap Timnas U-23 Indonesia pada Piala Asia AFC U-23 2024. Persepsi masyarakat terhadap Timnas Indonesia U-23 tidak hanya mempengaruhi dukungan selama turnamen berlangsung, tetapi juga citra timnas dan olahraga secara keseluruhan. Oleh karena itu, pemahaman yang baik tentang sentimen publik terhadap Timnas Indonesia U-23 dapat memberikan wawasan berharga bagi para pengambil keputusan dalam manajemen olahraga dan komunikasi media.

Tabel 1. *Review PICCO*

ANALISIS SENTIMEN PERSEPSI MASYARAKAT TERHADAP TIMNAS INDONESIA U-23 DALAM AFC-23 ASIAN CUP 2024 PADA MEDIA SOSIAL X MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

	Analisis sentimen pada media sosial X menggunakan metode naïve bayes.
Population	
Intervention	Bagaimana cara mendapatkan dan mengolah data tweet dari sosial media X agar dapat digunakan untuk analisis sentimen.

ANALISIS SENTIMEN PERSEPSI MASYARAKAT TERHADAP TIMNAS INDONESIA U-23 DALAM AFC-23 ASIAN CUP 2024 PADA MEDIA SOSIAL X MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Comparison	
Outcomes	Klasifikasi sentimen positif dan negative, dan Tingkat akurasi dari permodelan yang digunakan.
Context	Data berupa dataset publik yang didapat melalui crawling data X dengan google colab.

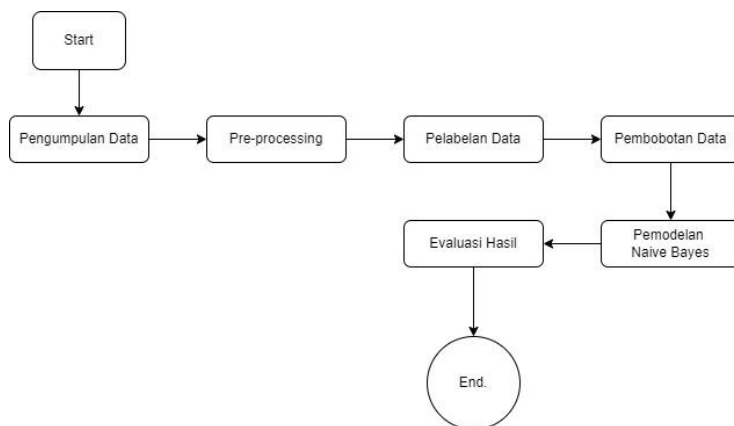
2. Metode Penelitian

Menurut Suharsimi Arikunto (2002), data penelitian mencakup semua bentuk fakta atau angka yang dapat dijadikan sumber untuk menyusun informasi yang relevan. Data ini, yang dikumpulkan dari berbagai sumber terpercaya, digunakan untuk menganalisis dan menjawab pertanyaan penelitian, menguji hipotesis, atau mencapai tujuan lain yang ditetapkan dalam penelitian ini. Penggunaan data yang valid dan dapat dipertanggungjawabkan adalah esensial untuk memastikan bahwa hasil penelitian dapat diandalkan dan relevan dengan tujuan penelitian. Analisis Sentimen adalah proses yang digunakan untuk mengidentifikasi dan mengekstrak informasi sentimen dari teks, yang biasanya dilakukan melalui penerapan *natural language processing (NLP)* dan teknik *machine learning*. Teknik-teknik ini bertujuan untuk mengklasifikasikan teks berdasarkan polaritas sentimen, apakah itu positif, negatif, atau netral. Salah satu metode yang sering digunakan dalam analisis sentimen adalah *sentiment classification*, yang memungkinkan peneliti untuk secara sistematis mengkategorikan teks ke dalam sentimen yang relevan berdasarkan konten dan konteks yang dianalisis.

Text Mining adalah pendekatan yang terdiri dari berbagai langkah untuk mengekstraksi informasi dari data teks yang tidak terstruktur. Proses *text mining* melibatkan serangkaian teknik yang dirancang untuk mengidentifikasi pola dalam data teks, memungkinkan klasifikasi dokumen berdasarkan topik yang diangkat dalam teks tersebut. *Text mining* sangat penting dalam mengolah data teks yang besar dan kompleks, memberikan kerangka kerja untuk analisis yang lebih mendalam dan memungkinkan penarikan kesimpulan yang relevan dari data yang dianalisis. Sepak bola, yang dikenal sebagai salah satu olahraga paling populer di dunia, adalah permainan yang dimainkan oleh dua tim dengan tujuan mencetak gol ke gawang lawan. Permainan ini melibatkan komposisi tim yang terdiri dari sebelas pemain di lapangan, dan memerlukan kombinasi pergerakan fisik, mental, serta koordinasi motorik halus dan kasar. Sepak bola juga menekankan pentingnya kerja sama tim yang solid untuk mencapai kemenangan.

Twitter, atau yang sekarang dikenal sebagai *X*, adalah salah satu platform media sosial yang paling populer di kalangan pengguna internet. Pengguna *Twitter* dapat menyebarkan postingan pendek yang dikenal sebagai *tweets*, yang dapat berisi teks, video, foto, atau tautan. Pada Oktober 2022, *Twitter* diakuisisi oleh Elon Musk, yang kemudian mengubah nama dan identitas platform ini menjadi *X*, termasuk mengganti logo burung biru yang ikonik dengan identitas baru yang lebih sesuai dengan visinya. Metode *Naïve Bayes* adalah pendekatan dalam *machine learning* yang menggunakan prinsip probabilitas untuk membangun model prediksi klasifikasi. Keunggulan metode ini terletak pada kemampuannya yang efektif dan cepat dalam mengolah data yang berjumlah besar, dengan memanfaatkan data kejadian masa lalu sebagai basis untuk prediksi. *Naïve Bayes* sering digunakan dalam analisis teks karena kemampuannya dalam mengklasifikasikan data dengan tingkat akurasi yang tinggi, meskipun asumsi yang digunakan relatif sederhana.

Penerapan Metodologi dalam penelitian ini dilakukan melalui beberapa tahapan yang telah dirancang secara sistematis untuk memastikan bahwa data yang diperoleh dianalisis dengan cara yang paling efektif. Setiap tahapan dalam metodologi ini dirancang untuk memaksimalkan validitas dan reliabilitas hasil yang diperoleh, dengan fokus utama pada pengumpulan data, pembersihan data, analisis data, dan interpretasi hasil.

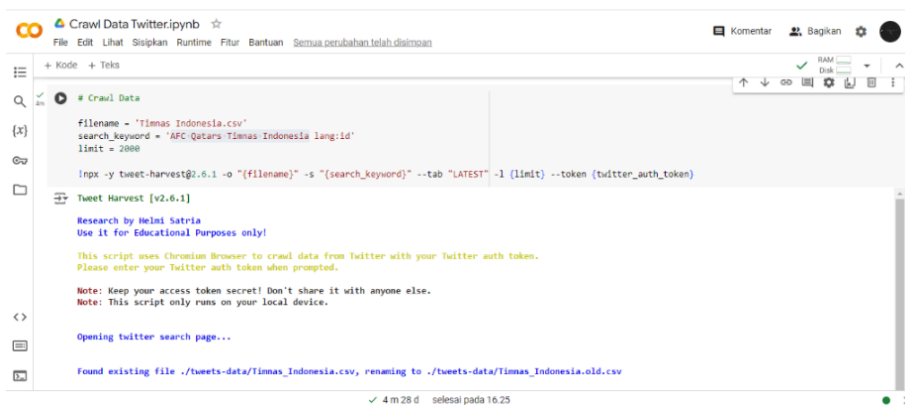


Gambar 1. Alur Penelitian

Dari Gambar 1 Alur Penelitian diatas dapat dijelaskan bahwa untuk melakukan penelitian analisis sentimen ini akan dilakukan dalam beberapa langkah yang terdiri pengumpulan dataset dengan menggunakan teknik crawling data X menggunakan google colab. Data yang telah di ambil dari X akan dilakukan pelabelan, lalu dilanjutkan dengan tahap preprocessing data seperti case folding, filtering, stopward removal, dan tokenize, setelah itu dilakukan pembobotan data dengan metode TF-IDF, lalu proses klasifikasi dengan metode Naïve Bayes dengan menggunakan 2 kategori data, yaitu data latih dan data uji, proses selanjutnya evaluasi hasil dengan menentukan nilai akurasi, precision yang diperoleh dari hasil klasifikasi untuk mendapatkan output yang diinginkan.

2.1 Pengumpulan Data

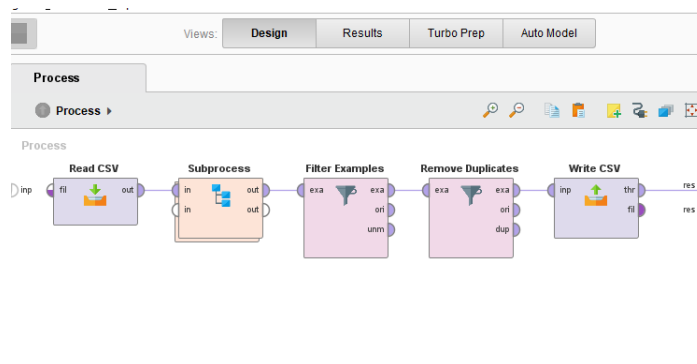
Pengumpulan data merupakan langkah awal dalam “Analisis Sentimen Persepsi Masyarakat Terhadap Timnas Indonesia U-23 Dalam Afc-23 Asian Cup 2024 Pada Media Sosial X Menggunakan Metode Naïve Bayes Classifier”. Data yang diambil berupa data tweet pada media social X menggunakan API Twitter dengan Teknik crawling data menggunakan Google Colab. Setelah data dikumpulkan, kemudian data disimpan dengan format CSV.



Gambar 2. Crawling Data Tweet

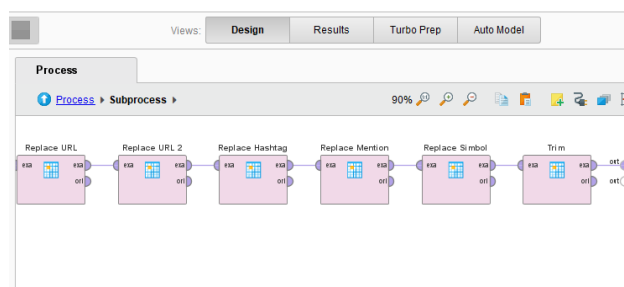
2.2 Cleaning Data

Selanjutnya data tweet yang telah dikumpulkan melalui crawling data akan memasuki proses cleaning data yang bertujuan untuk membersihkan Tweet dari kata-kata yang tidak diperlukan atau duplikat, lalu simbol-simbol seperti karakter hastag “#”, metion “@”, kemudian membuang link url dan angka yang tidak dibutuhkan dalam proses sentimen analisis.



Gambar 3. Operator Cleaning Data

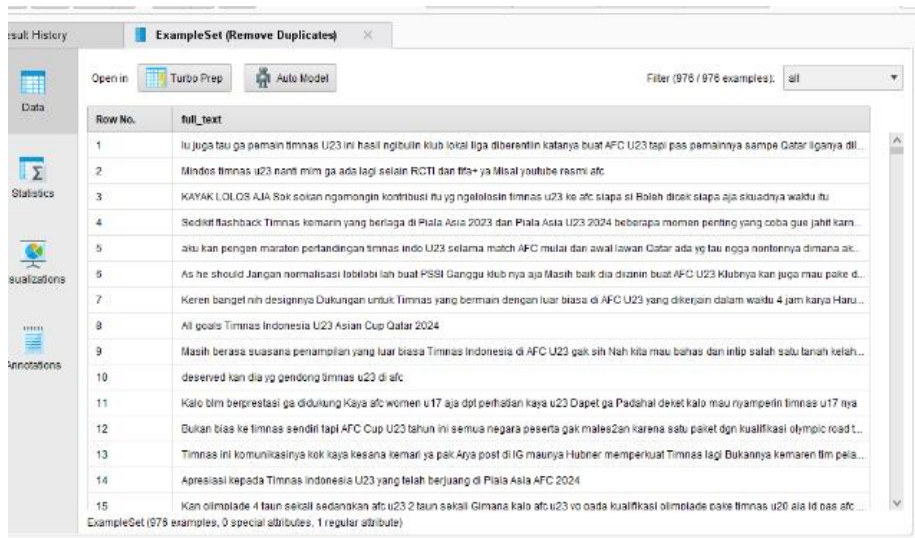
Proses *cleaning data* pada *tweet* menggunakan *RapidMiner Studio* dilakukan melalui beberapa tahapan penting. Pertama, data yang telah dikumpulkan sebelumnya diimpor menggunakan operator *Read CSV*, di mana dataset *tweet* disimpan dalam format CSV. Langkah ini memastikan bahwa data yang akan diproses tersedia dalam format yang mudah diakses dan diolah. Selanjutnya, proses ini dilanjutkan dengan penerapan *Subprocess* yang dirancang untuk meningkatkan keterbacaan dan pemeliharaan alur proses. *Subprocess* ini berfungsi agar alur kerja menjadi lebih efisien dan mempermudah pembaca dalam memahami struktur dan tujuan dari setiap bagian proses. Untuk memastikan analisis yang lebih fokus dan relevan, operator *Filter Examples* digunakan untuk membantu pengguna memfokuskan analisis pada data yang diperlukan dan menghilangkan data yang tidak relevan atau diperlukan. Kemudian, operator *Remove Duplicates* diterapkan untuk memastikan bahwa setiap contoh dalam dataset adalah unik, yang merupakan langkah penting dalam banyak aplikasi analisis data untuk menghindari redundansi dan memastikan akurasi hasil analisis. Setelah proses *cleaning* selesai, data yang telah diproses disimpan kembali menggunakan operator *Write CSV*. Langkah ini memungkinkan data yang telah dibersihkan untuk disimpan dalam format CSV, sehingga mudah dibaca dan dapat diolah lebih lanjut oleh berbagai perangkat lunak lain yang mendukung format ini.



Gambar 4. Subprocess Cleaning Data Menggunakan Operator Replace

Penulis akan menggunakan beberapa operator *Subprocess* seperti *Replace* dan *Trim* untuk membersihkan *tweet* dari kata-kata yang tidak diperlukan. Pertama, operator *Select Attributes* digunakan untuk memilih satu atribut dalam dataset yang akan dibersihkan, dengan parameter yang ditetapkan pada atribut *full_text*. Selanjutnya, operator *Replace RT* berfungsi untuk menghapus kata "RT" dalam

dataset, sehingga memastikan bahwa retweet tidak disertakan dalam analisis. Operator *Replace URL* dan *Replace URL 2* digunakan untuk menghapus tautan URL yang muncul di awal, tengah, atau akhir kalimat, dengan parameter yang disesuaikan untuk mencakup semua posisi URL dalam teks. Selain itu, simbol *hashtag* dan *mention* dihapus menggunakan operator *Replace Hashtag* dan *Replace Mention*, dengan parameter yang dirancang untuk menghapus simbol-simbol tersebut baik di awal maupun di tengah kalimat. Operator *Replace Simbol* juga diterapkan untuk menghapus berbagai simbol yang ada dalam dataset, termasuk karakter-karakter seperti tanda seru, tanda tanya, koma, titik koma, dan lainnya. Terakhir, operator *Trim* digunakan untuk membersihkan data teks dengan menghapus karakter kosong atau *whitespace* yang terdapat di awal dan akhir string dalam atribut teks, sehingga menghasilkan data yang lebih bersih dan siap untuk dianalisis lebih lanjut.



Gambar 5. Tweet Dataset Cleaning Results

Dari total 1200 dataset tweet mentah yang belum diolah, setelah melalui tahapan cleaning data menggunakan RapidMiner Studio, kini tersisa 973 dataset tweet yang telah dibersihkan. Proses cleaning data ini melibatkan beberapa langkah penting, termasuk penghapusan entri duplikat, pemangkasan beberapa atribut yang tidak diperlukan, serta penyaringan tweet berdasarkan kriteria tertentu. Dengan demikian, dataset yang dihasilkan lebih bersih dan siap untuk dilakukan analisis sentimen lebih lanjut.

2.3 Pelabelan Data

Dalam tahapan pelabelan data telah dilakukan pelabelan manual sebanyak 972 data tweet. Data tersebut dibagi menjadi dua kategori. Yaitu label *“Positive”* dan label *“Negative”*

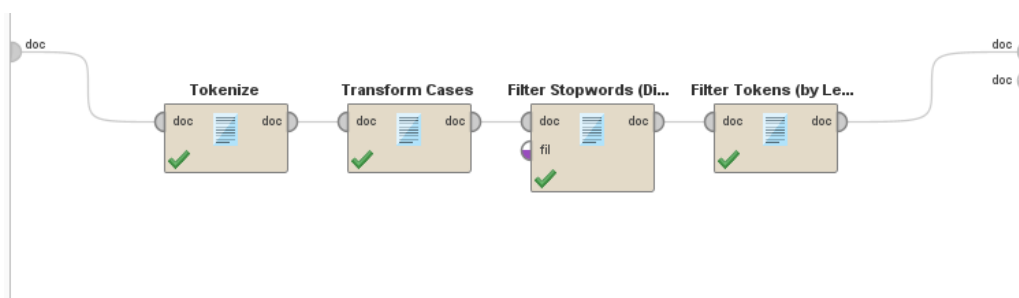
Tabel 2. Pelabelan Data

No	Text	Sentiment
1	lu juga tau ga pemain timnas U23 ini hasil ngibulin klub lokal liga dibereentiin katanya buat AFC U23 tapi pas pemainnya sampe Qatar liganya dilanjut Klub lokal ga lepas pemain diserang netizen + federasi goblokmj itu padahal bukan agenda FIFA giliran klub luar dimaklumin	Negative
2	Mindos timnas u23 nanti mlm ga ada lagi selain RCTI dan fifa+ ya Misal youtube resmi afc	Positive
3	KAYAK LOLOS AJA Sok sokan ngomongin kontribusi itu yg ngelolosin timnas u23 ke afc siapa si Boleh dicek siapa aja skuadnya waktu itu	Negative

4	Sedikit flashback Timnas kemarin yang berlaga di Piala Asia 2023 dan Piala Asia U23 2024 beberapa momen penting yang coba gue jahit karna vibes nyai? AFC Asian Cup	Negative
5	aku kan pengen maraton pertandingan timnas indo U23 selama match AFC mulai dari awal lawan Qatar ada yg tau ngga nontonnya dimana aku cari di youtube gak nemu full matchnya, kebanyakan cuma highlight aja, kalo punya linknya kasitau ya	Positive

2.4 Preprocessing Data

Tahap *Preprocessing data* pada penelitian ini adalah untuk mengubah data mentah menjadi data yang siap diolah.



Gambar 6. *Process Preprocessing data*

Tahapan *preprocessing* data yang dilakukan melibatkan beberapa langkah penting untuk mempersiapkan data sebelum analisis lebih lanjut. Pertama, proses *Tokenize* digunakan untuk memecah teks menjadi unit-unit kecil yang disebut *token*, sehingga setiap kata atau frasa dalam teks dapat dianalisis secara individual. Setelah itu, langkah *Transform Cases* diterapkan untuk mengubah semua huruf dalam teks menjadi huruf kecil, yang bertujuan untuk menyamakan format teks dan menghindari perbedaan yang disebabkan oleh penggunaan huruf besar dan kecil. Selanjutnya, *Filter Stopword (Dictionary)* digunakan untuk menghapus kata-kata yang tidak memiliki arti penting dalam analisis teks, seperti "dan", "atau", "tetapi", dan kata-kata lain yang tidak memberikan kontribusi signifikan terhadap pemahaman konteks teks. Terakhir, langkah *Filter Token (By Length)* diterapkan untuk memfilter *token* berdasarkan panjangnya, yang memungkinkan penghapusan kata atau istilah yang terlalu pendek atau terlalu panjang untuk relevan dalam analisis. Tahapan ini memastikan bahwa data yang digunakan dalam analisis adalah data yang bersih dan relevan, siap untuk dianalisis lebih lanjut.

2.5 Pembobotan Data

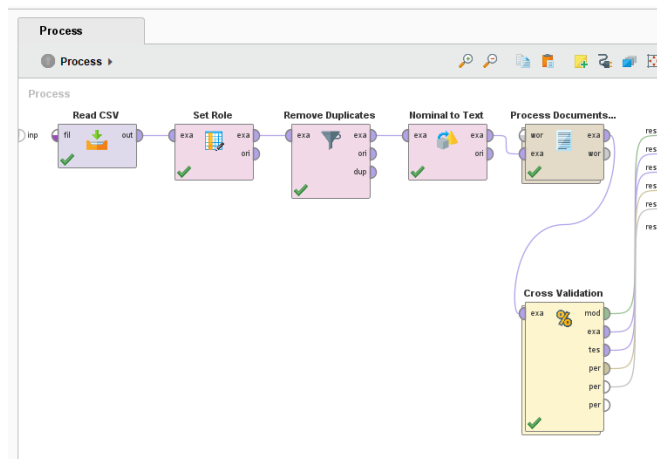
Pembobotan kata di dapat dari tahapan preprocessing data yang telah dilakukan pada tahap sebelumnya. Metode yang digunakan penulis dalam pembobotan kata adalah algoritma TF-IDF (*Term Frequency-Inverse Document Frequency*), TF-IDF menggabungkan dua metrik: *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). Dengan menggunakan TF-IDF, penulis dapat mengidentifikasi kata-kata yang paling relevan dan penting dalam dataset, dan membantu tugas-tugas dalam melakukan klasifikasi data.

Row No.	Sentimen	text
1	Positive	keren banget desainnya dukungan timnas bermain dikerjain karya dibikin
2	Positive	goals timnas indonesia asian qatar
3	Negative	viralkan ronaldo messi salto kelucuan wasit thailand timnas indonesia kalah uzbekistan piala dunia
4	Negative	olahraga bawa nama negara team fomo ngidol ngebiasin idol ngikutin timnas kmm idolizing player bibit awas smpe
5	Positive	olympic bernasib timnas indonesia juara anaknya manis kalem ulet tenaganya palibg gila seantero jarang lelah kecuai cidera
6	Positive	soroti pernyataan shin yong laga menegangkan lawan irak perebutan juara peluang lolos olimpiade paris kalah menyakitkan irak pela...
7	Positive	bangga timnas muda semoga kedepan juara
8	Positive	kemeriahan pendukung timnas indonesia nobar indonesia iraq bale kota tasikmalaya
9	Positive	setuju netjen fomo nonton timnas semenjak semifinal marselino maen pertandingan bagus
10	Positive	timnas sepakbola masuk semifinal uber masuk semifinal waaaaahhh proud
11	Negative	damn liat kemaren bener digendong pemain naturalisasi banget permainan timnas fakiu towel lokalpret anjing
12	Positive	terimakasih timnas team berjuang https ucpxkzodfw
13	Negative	rekap netjen indonesia kalah lawan qatar wasit serang menang lawan ausi emando puja menang lawan jordan lolos final indonesia L...

Gambar 7. Hasil Pembobotan Data

2.6 Pemodelan *Naïve Bayes*

Workflow dari tahapan ini akan menghasilkan output sebuah model klasifikasi dengan metode *Naïve Bayes* yang akan digunakan dalam proses sentimen analisis. Proses tahapan pembuatan model seperti pada Gambar 4.2.6. *Workflow* Pembuatan *Model Naïve Bayes*.



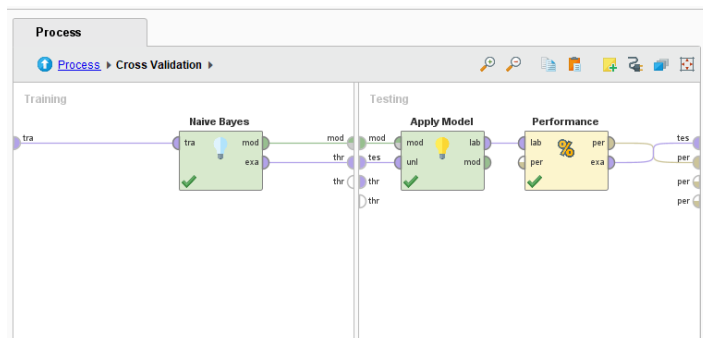
Gambar 8. *Workflow* Pembuatan Model *Naïve Bayes*

3. Hasil dan Pembahasan

3.1 Hasil

3.1.1 Pengujian

Tahapan terakhir pada penelitian ini adalah tahap pengujian. Dari proses pembuatan model naïve bayes sebelumnya akan dilakukan penerapan model untuk melakukan prediksi pada sentimen dataset.



Gambar 9. Pengujian Model Naïve Bayes

3.1.2 Hasil Akurasi

Hasil perhitungan akurasi dataset menggunakan metode *Naïve Bayes*, didapatkan nilai *Accuracy* sebesar 73,94%. Penulis melabeli data sebagai sentimen positif dan sentimen negatif secara manual.

Tabel 3. Nilai Akurasi

	<i>Accuracy: 73,94% +/- 3.45% (micro average: 73,94%)</i>		
	<i>True Negative</i>	<i>True Positive</i>	<i>Class Precision</i>
<i>Pred. Negative</i>	43	153	21,94%
<i>Pred. Positive</i>	100	675	87,10%
<i>Class Recall</i>	30,07%	81,52%	

3.1.3 Hasil Pengujian

Hasil akhir dari metode pengujian ini, yaitu hasil prediksi Sentimen Terhadap Timnas Indonesia U-23 Dalam Afc-23 Asian Cup 2024, berdasarkan data yang didapat dari Sosial media X dan diimplementasikan dengan metode *Naïve Bayes* menunjukkan nilai *Accuracy* sebesar 73,94% nilai *Recall* Positif sebesar 81,51%, nilai *Recall* Negatif sebesar 0%, nilai *Precision* Positif sebesar 87.09%, nilai *Precision* Negatif sebesar 0%. Dari 972 dataset, terdapat 143 data terprediksi negatif dan 828 data yang terprediksi Positif.

```

PerformanceVector
PerformanceVector:
accuracy: 73.94% +/- 3.45% (micro average: 73.94%)
ConfusionMatrix:
True:  Negative      Positive
Negative:    43       153
Positive:   100       675
AUC (optimistic): 0.865 +/- 0.028 (micro average: 0.865) (positive class: Positive)
AUC: 0.500 +/- 0.000 (micro average: 0.500) (positive class: Positive)
AUC (pessimistic): 0.248 +/- 0.073 (micro average: 0.248) (positive class: Positive)
precision: 87.09% +/- 1.44% (micro average: 87.10%) (positive class: Positive)
ConfusionMatrix:
True:  Negative      Positive
Negative:    43       153
Positive:   100       675
recall: 81.51% +/- 3.76% (micro average: 81.52%) (positive class: Positive)
ConfusionMatrix:
True:  Negative      Positive
Negative:    43       153
Positive:   100       675
    
```

Gambar 10. Hasil Pengujian Model Naïve Bayes

3.2 Pembahasan

Penelitian ini mengadopsi metodologi yang cermat dan sistematis untuk menganalisis sentimen masyarakat terhadap Timnas Indonesia U-23 selama kompetisi AFC U-23 Asian Cup 2024, dengan memanfaatkan media sosial X (sebelumnya dikenal sebagai *Twitter*) sebagai sumber data utama. Proses *preprocessing* dan *cleaning* data sangat penting dalam memastikan kualitas dan keakuratan analisis yang dilakukan. Tahapan *preprocessing* yang meliputi *tokenization*, *transform cases*, *filter stopwords*, dan *filter tokens by length* memungkinkan peneliti untuk menyaring data dan fokus pada elemen-elemen yang paling

relevan, seperti yang telah dikemukakan dalam penelitian oleh Fairuz *et al.* (2021) yang menggunakan metode serupa untuk analisis sentimen terhadap COVID-19 di *Twitter*. Pentingnya penggunaan teknik *text mining* dalam memproses data teks yang tidak terstruktur juga ditegaskan dalam penelitian ini. Sebagaimana dijelaskan oleh Kusuma dan Nugroho (2021), *text mining* memungkinkan peneliti untuk mengidentifikasi pola dalam data dan mengklasifikasikan dokumen berdasarkan topik yang relevan. Dalam konteks penelitian ini, *text mining* membantu dalam mengorganisir dan menyaring data *tweet* yang besar dan kompleks, yang kemudian diproses lebih lanjut menggunakan metode *Naïve Bayes*.

Metode *Naïve Bayes* yang digunakan dalam penelitian ini telah terbukti efektif dalam berbagai studi sebelumnya untuk mengklasifikasikan sentimen dalam data teks, seperti yang juga ditemukan oleh Maria *et al.* (2023) dalam analisis sentimen terhadap penggunaan aplikasi MyPertamina di *Twitter*. Metode ini menawarkan keuntungan dalam kecepatan dan efisiensi, terutama ketika berhadapan dengan data dalam jumlah besar. Dalam penelitian ini, *Naïve Bayes* digunakan untuk mengklasifikasikan sentimen *tweet* sebagai positif atau negatif, dengan hasil yang menunjukkan bahwa mayoritas sentimen masyarakat terhadap Timnas Indonesia U-23 adalah positif. Temuan ini sejalan dengan studi sebelumnya oleh Sulastiyono *et al.* (2023), yang juga menunjukkan efektivitas *Naïve Bayes* dalam mengklasifikasikan sentimen terhadap isu-isu publik lainnya di Indonesia. Hasil analisis menunjukkan bahwa metode *Naïve Bayes* mampu mengelola kompleksitas data dengan baik, memberikan gambaran yang akurat mengenai persepsi masyarakat terhadap kinerja Timnas Indonesia U-23. Sebagai contoh, penelitian oleh Matheos Sarimole dan Kudrat (2024) yang mengkaji sentimen terhadap aplikasi Satu Sehat di *Twitter* juga mengonfirmasi bahwa *Naïve Bayes* adalah alat yang efektif untuk menganalisis data sentimen dengan akurasi yang signifikan.

4. Kesimpulan

Penelitian ini menggunakan metode *Naïve Bayes* untuk mengklasifikasikan sentimen publik terhadap Timnas Sepakbola Indonesia U-23 selama turnamen di evaluasi menggunakan Software RapidMiner Studio. Berdasarkan nilai akurasi hasil klasifikasi sebesar 73,94% dari 972 dataset komentar yang di analisis, dan hasil klasifikasi total data, 828 komentar yang di klasifikasikan sebagai positif dan 143 sebagai negative. Menunjukkan bahwa metode *Naïve Bayes* cukup efektif dalam menganalisis sentimen publik di media sosial, memberikan gambaran yang akurat mengenai persepsi masyarakat terhadap kinerja Timnas Sepakbola Indonesia U-23 selama turnamen tersebut.

5. Ucapan Terima Kasih

Peneliti mengucapkan terima kasih yang sebesar-besarnya atas bimbingan, dukungan, dan dedikasi Bapak Untung Surapati, SE., SH., MH., MM., M. Kom dalam membantu pembuatan jurnal ini. Tanpa saran, pengarahan dan motivasi yang berharga dari Anda, pencapaian ini tidak akan mungkin terwujud. Saya sangat berterima kasih atas waktu dan pengetahuan yang Anda berikan untuk membimbing saya dalam menyusun jurnal ini.

6. Daftar Pustaka

- Akbar, Y., & Ihsan, A. N. (2023). Analisis Sentimen Twitter Terhadap Opini Masyarakat Pada Sea Games Kamboja 2023 Menggunakan Algoritma Support Vector Machine. *INTECOMS: Journal of Information Technology and Computer Science*, 6(2), 814-821. DOI: <https://doi.org/10.31539/intecom.v6i2.7670>.
- Aufa, M. M. (2021). Instrumen tes menggiring bola pada permainan sepakbola. *Jurnal Edukasimu*, 1(2).
- Chairunnisa, C. (2017). Pengaruh Literasi Membaca Dengan Pemahaman Bacaan (Penelitian Survei pada Mahasiswa STKIP Kusumanegara Jakarta). *Jurnal Tukuran*, 6(1), 745-756. DOI: <https://doi.org/10.33603/jt.v6i1.1584>.
- Fairuz, A. L., Ramadhani, R. D., & Tanjung, N. A. F. (2021). Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Journal of Dinda: Data Science, Information Technology, and Data Analytics*, 1(1), 42-51.
- Islamy, I. (2019). Penelitian Survei dalam Pembelajaran dan Pengajaran Bahasa Inggris. *Purwokerto: Universitas Muhammadiyah Purwokerto*.
- Kusuma, A., & Nugroho, A. (2021). Analisa Sentimen Pada Twitter Terhadap Kenaikan Tarif Dasar Listrik Dengan Metode Naïve Bayes. *Jurnal Ilmiah Teknologi Informasi Asia*, 15(2), 137-146. DOI: <https://doi.org/10.32815/jitika.v15i2.557>.
- Maria, R., Umayah, R. U., Mahardinny, S., Kalana, D., & Saputra, D. D. (2023). Analisis Sentimen Persepsi Masyarakat Terhadap Penggunaan Aplikasi My Pertamina Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier. *Jurnal Komputer Antartika*, 1(1), 1-10. DOI: <https://doi.org/10.70052/jka.v1i1.1>.
- Safira, A., & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier. *ZONAsi: Jurnal Sistem Informasi*, 5(1), 59-70. DOI: <https://doi.org/10.31849/zn.v5i1.12856>.
- Sarimole, F. M., & Kudrat, K. (2024). Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine. *Jurnal Sains dan Teknologi*, 5(3), 783-790. DOI: <https://doi.org/10.55338/saintek.v5i3.2702>.
- Sulastiyono, R., Setiawan, A., & Nugroho, S. (2023). Sentimen Analisis Pembatalan Indonesia Menjadi Tuan Rumah Piala Dunia U-20 Menggunakan Metode Naïve Bayes. *Journal of Information System Research (JOSH)*, 4(4), 1387-1394. DOI: <https://doi.org/10.47065/josh.v4i4.3737>.
- Yessy Asri, S. T., Kuswardani, D., & Kom, M. (2024). *MACHINE LEARNING & DEEP LEARNING: Analisis Sentimen Menggunakan Ulasan Pengguna Aplikasi*. Uwais Inspirasi indonesia.