

ANALISIS SENTIMEN TERHADAP WARGA NET PADA LAGA AFC U23 INDONESIA VS KOREA SELATAN

M. Rifqi Mikatama^{1a,*}, Muhammad Reza Redo Islami^{2b}

^{1,2}Program Studi Teknik Informatika, STMIK Dharmawacana
Metro, Lampung Indonesia

¹rifqimikatama@gmail.com
²reza.redo@hotmail.com

Abstract

This research aims to analyze netizens' sentiments towards the football match between the Indonesian U23 national team and South Korea in the AFC U23 Asian Cup. The research method used is sentiment analysis on data obtained from Twitter social media. The analysis process includes data cleaning, text preprocessing, and labeling positive, negative, and neutral sentiment. The research results show that in general netizens' sentiment towards the match tends to be positive. In fact, netizens expressed more support and appreciation for the two teams than negative expressions. This research provides important insights for decision makers such as football clubs and event organizers to adapt communications strategies, increase fan engagement and manage reputation more effectively based on netizen sentiment. From this research, it was concluded that this match was able to trigger enthusiasm and nationalism from netizens.

Keywords : Sentiment Analysis, Afc U23, Twitter, Netizen

Abstrak

Penelitian ini bertujuan untuk menganalisis sentimen warga net terhadap pertandingan sepak bola antara timnas U23 Indonesia dan Korea Selatan dalam ajang Piala Asia AFC U23. Metode penelitian yang digunakan adalah analisis sentimen pada data yang diperoleh dari media sosial twitter. Proses analisis melibatkan pembersihan data, preprocessing teks, dan pelabelan sentimen positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa secara umum sentimen warga net terhadap pertandingan tersebut cenderung positif. Faktanya, warga net lebih banyak menyuarakan dukungan dan apresiasi kepada kedua tim dibandingkan dengan ekspresi negatif. Penelitian ini memberikan wawasan penting bagi pengambil keputusan seperti klub sepak bola dan penyelenggara acara untuk menyesuaikan strategi komunikasi, meningkatkan keterlibatan penggemar, dan mengelola reputasi lebih efektif berdasarkan sentimen warga net. Dari penelitian ini menyimpulkan bahwa pertandingan tersebut mampu memicu antusiasme dan nasionalisme dari warga net.

Keywords : Analisis Sentimen, Afc U23, Twitter, Warga Net

1. PENDAHULUAN

Perkembangan media sosial saat ini sangat pesat, telah banyak platform media sosial yang popular salah satunya adalah Twitter. Seperti media sosial lainnya, Twitter adalah platform media sosial yang memungkinkan pengguna untuk berinteraksi secara pribadi atau publik. Berbagai topik dibicarakan di Twitter, antara lain politik, ekonomi, sosial, budaya dan hukum(Aryanti, 2022). Analisis sentimen biasa digunakan untuk mengukur tingkat kecenderungan seseorang terhadap suatu hal dalam sebuah pembicaraan. Penggalian informasi dalam opini dapat menjadi sebuah studi untuk mencari kecenderungan seseorang dan dapat digunakan dalam berbagai macam bidang (Abdulloh & Pambudi, 2021).

Salah satu trending topik terkait dengan isu laga AFC U23 pada tanggal 25-April_2024 Indonesia melawan Korea Selatan. Yang menjadikannya trending topic warga net ialah Indonesia yang berhasil mengalahkan Korea Selatan pada perempat final. Pada kompetisi olahraga warga net menjadi sangat aktif dalam mengungkapkan sentimen mereka terkait pertandingan dan hasilnya melalui media sosial seperti Twitter, Instagram, dan Facebook (Tamami, n.d.)

Dalam penelitian ini diperoleh data dari twitter dengan jumlah *rows* awal sebanyak 5994 dan 15 kolom. Kolom yang digunakan dalam penelitian hanya 3 kolom yaitu *kolom full_text*, *created_at*, dan *username*. Dalam penelitian ini *kolom full_text* akan dijadikan target utama dalam penelitian mencari sentimen warga net. Penelitian ini bertujuan untuk menganalisis sentimen warga net terhadap pertandingan sepak bola antara timnas U23 Indonesia dan Korea

Selatan dalam ajang Piala Asia AFC U23. Analisis sentimen diharapkan dapat memberikan gambaran tentang bagaimana warga net menyikapi pertandingan tersebut. Data yang diperoleh melalui twitter yang memiliki jumlah pengguna sebesar 63,6% dari populasi penduduk Indonesia dapat memberikan gambaran tentang sentimen masyarakat terhadap pertandingan AFC U23 antara Indonesia dan Korea Selatan (Permana et al., 2021).

Salah satu pendekatan yang dapat digunakan untuk menganalisis sentimen warga net terhadap suatu peristiwa adalah Sentiment Analysis atau analisis sentimen (Putra et al., 2022). Metode ini melibatkan pengambilan data teks yang diunggah oleh individu ke media sosial dan menganalisisnya untuk mengetahui pandangan, opini, dan perasaan mereka terkait suatu peristiwa atau topik tertentu. Dalam melakukan analisis sentimen, kita dapat memanfaatkan data teks yang diunggah oleh warga net yang mengomentari pertandingan tersebut.

Cleansing data memainkan peran penting dalam siklus hidup data. Masalah kualitas data dapat menyebabkan informasi yang diekstraksi menjadi terdistorsi atau tidak dapat digunakan sebuah masalah yang dapat dikurangi atau dihilangkan melalui pembersihan data. Beberapa masalah dapat dicegah selama pengumpulan data, namun banyak pula yang harus ditangani pada tahap pembersihan data (Huang et al., 2019). Dalam penelitian ini dilakukan *cleaning data duplicated, null, mention, hastagh, retwet, url, non-alfanumerik, dan spasi* berlebihan.

Preprocessing data text dalam bahasa Indonesia merujuk pada serangkaian langkah untuk menormalkan, menghilangkan, dan mempersiapkan. Pada *preprocessing data text* tahap ini adalah tahapan di mana data disiapkan agar menjadi data yang siap untuk dianalisis (Hadna et al., 2016). Pada penelitian ini tahapan *preprocessing* meliputi proses *stopwords, tokenizing, dan stemming*. Dalam penelitian ini melakukan aktifitas *translate data* ke dalam bahasa indonesia disebabkan saat proses crawling data, data yang diambil mencakup keseluruhan warga net yang turut andil dalam menanggapi pertandingan AFC U23 Indonesia vs Korea Selatan. Dalam analisis sentimen terhadap warga net pada pertandingan AFC U23 Indonesia vs Korea Selatan, penelitian ini menggunakan metode analisis sentimen pada data yang telah dilabeli. Metode analisis sentimen adalah metode yang digunakan untuk mengidentifikasi dan menganalisis sentimen atau opini dalam teks berdasarkan polaritasnya, yaitu apakah sentimen tersebut positif, negatif, atau netral (Sari et al., 2021).

Penelitian ini berpotensi memberikan wawasan berbasis data yang penting bagi para pengambil keputusan. Misalnya klub sepak bola, penyelenggara acara, dan stakeholder terkait lainnya dapat menggunakan hasil analisis ini untuk menyesuaikan strategi komunikasi mereka, meningkatkan keterlibatan penggemar, dan mengelola reputasi dengan lebih efektif. Pada penelitian ini untuk menampilkan hasil akhir analisis sentimen menggunakan visualisasi data yang dapat membantu menampilkan hasil sentimen. Hasil visualisasi data yang telah dilabeli dapat divisualisasikan dalam bentuk grafik atau diagram untuk memperoleh pemahaman yang lebih jelas tentang persebaran sentimen pada pertandingan AFC U23 Indonesia vs Korea Selatan (Hanafi & Solichin, 2023). Hasil dari analisis sentimen terhadap warga net pada pertandingan AFC U23 Indonesia vs Korea Selatan menunjukkan memiliki beragam sentimen yang muncul dari warga net.

2. KERANGKA TEORI

2.1 Media Sosial

Media sosial merupakan platform digital yang memungkinkan penggunaannya untuk berinteraksi dan berbagi konten (Azzahra, 2021). Melalui media sosial, pengguna dapat dengan mudah menyuarakan pendapat, melakukan diskusi, dan bertukar informasi tanpa ada batasan ruang dan waktu. Menurut (Kaplan & Haenlein, 2010), media sosial memiliki peran penting dalam menyebarkan informasi dan membentuk opini publik. Dalam konteks olahraga, media sosial memungkinkan penggemar untuk berpartisipasi aktif dalam diskusi dan mendukung tim favorit mereka.

2.2 Pengaruh Media Sosial dalam Olahraga

(Sanderson, 2011) menyatakan bahwa media sosial telah mengubah cara penggemar berinteraksi dengan olahraga. Platform seperti Twitter dan Facebook memungkinkan penggemar untuk memberikan tanggapan *real-time* selama pertandingan, berbagi momen penting, dan menyatakan dukungan atau ketidakpuasan mereka terhadap hasil pertandingan.

2.3 Analisis Sentimen

Analisis sentimen merupakan proses pengumpulan, pemrosesan, dan evaluasi pendapat, opini, atau sentimen dari teks yang diekspresikan oleh individu atau kelompok. Tujuan utama dari analisis sentimen adalah untuk memahami sikap, emosi, dan persepsi orang terhadap suatu topik atau entitas tertentu. Dalam analisis sentimen perlu dilakukannya pra-pemrosesan sentimen, tujuan dilakukannya hal ini guna memperoleh hasil sentimen yang baik, berikut step by step pra-pemrosesan nya :

2.3.1 pemrosesan data

Pra-pemrosesan data melibatkan langkah-langkah seperti pembersihan data, tokenisasi, dan penghilangan stopwords untuk mempersiapkan teks sebelum analisis sentimen dilakukan. (Aggarwal & Zhai, 2012) menekankan pentingnya langkah-langkah ini dalam memastikan akurasi analisis sentimen.

2.3.2 Teknik Analisis Sentimen

Teknik analisis sentimen melibatkan berbagai metode dan alat untuk mengidentifikasi dan mengukur sentimen dalam teks. Beberapa alat yang umum digunakan termasuk pustaka seperti NLTK (Natural Language Toolkit) menyediakan berbagai alat untuk pemrosesan bahasa alami, termasuk tokenisasi, stemming, dan analisis sentimen dasar. TextBlob adalah pustaka yang dibangun di atas NLTK dan menyediakan antarmuka yang lebih sederhana untuk analisis sentimen, memungkinkan pengguna untuk dengan mudah mendapatkan sentimen positif, negatif, dan netral dari teks.

2.3.3 Kategorisasi Sentimen

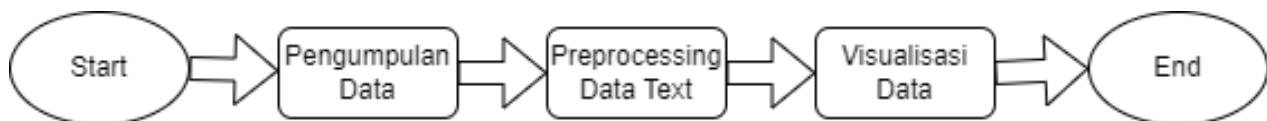
Pendekatan untuk mengkategorikan sentimen biasanya melibatkan klasifikasi teks menjadi tiga kategori utama: positif, negatif, dan netral. (Pang & Lee, 2008) menjelaskan metode dan teknik yang dapat digunakan untuk mengkategorikan sentimen dalam analisis teks. Kategorisasi ini sering dilakukan menggunakan algoritma pembelajaran mesin yang dilatih pada dataset dengan label sentimen yang jelas.

2.4 Faktor-faktor Mempengaruhi Sentiment

Faktor yang dapat mempengaruhi sentimen pada laga AFC U23 seperti performa tim, keputusan wasit, dan peristiwa penting dalam pertandingan dapat mempengaruhi hasil sentimen. Insiden-insiden kontroversial selama pertandingan dapat menjadi pemicu utama pada hasil sentimen negatif di media sosial. Interaksi dan diskusi antar pengguna dapat menimbulkan hasil sentimen tertentu dan dapat membentuk suatu opini tertentu.

3. METODOLOGI

Analisis sentimen atau *opinion mining* merupakan pengaplikasian dari ilmu *komputasi linguistik, text analysis, dan natural language processing* yang digunakan untuk mengekstrak subjektifitas dari sebuah sumber teks(Husada & Paramita, 2021). Dalam penelitian ini dilakukan beberapa tahapan yang saling berkaitan antara satu dengan tahap selanjutnya. Struktur tahapan penelitian yang dilakukan dapat dilihat pada gambar 1.



Gambar 1. Flowchart Analisis Sentimen

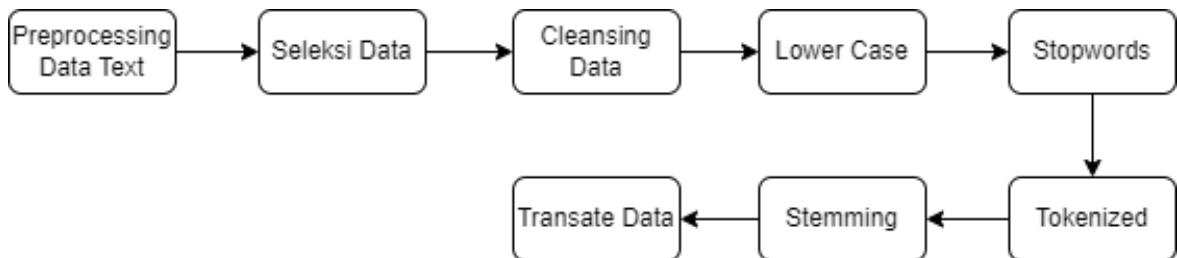
Pada penelitian ini, data yang diperoleh diolah menggunakan Python. Proses pengolahan data mencakup beberapa tahapan utama, yaitu pengumpulan data, *preprocessing data text*, dan visualisasi. Tahapan-tahapan tersebut dijelaskan secara rinci sebagai berikut:

3.1 Pengumpulan Data

Pengumpulan data adalah proses mengumpulkan informasi dari berbagai sumber untuk digunakan dalam analisis dan penelitian. Dalam konteks analisis sentimen, pengumpulan data merujuk pada langkah-langkah yang diambil untuk mendapatkan teks yang akan dianalisis sentimennya. Penelitian ini diperoleh melalui penerapan teknik crawling dengan menggunakan kata kunci "AFC_U23", "Indonesia_vs_Korea Selatan" pada platfrom sosial media Twitter atau X . Pengumpulan data dilakukan dengan memanfaatkan *teknik crawling* data, menggunakan *library* tweet harvest yang tersedia dalam bahasa pemrograman python. Data yang terkumpul sebanyak 5994 data. Dataset tersebut kemudian dilakukan *preprocessing*, yakni membersihkan data, menghapus karakter yang tidak perlu, pelabelan data(Ananda & Suryono, 2024). *Crawling* data merupakan langkah yang penting dalam penelitian analisis sentimen karena data twitter merupakan sumber informasi utama untuk menilai *opini public* (Ulfatriyani et al., 2020). Data yang diperoleh melalui Twitter atau X terdapat pada gambar data awal.

3.2 Preprocessing data text

Preprocessing data text merupakan langkah yang diambil untuk mempersiapkan suatu text yang berada pada suatu data mentah sebelum dianalisis lebih lanjut(Hudha et al., 2022). Pada proses ini digunakan python untuk melakukan *seleksi data, cleansing data, tokenized, stopwords, stemming*. Dalam penelitian ini, teknik *preprocessing data text* digunakan untuk mempersiapkan teks yang akan dianalisis lebih lanjut dalam analisis sentimen terhadap warga net pada pertandingan AFC U23 Indonesia vs Korea Selatan berikut struktur tahap *preprocessing data text* dapat dilihat pada gambar 2, serta penjelasan dari tahap *preprocessing data text*:



Gambar 2. Flowchart Preprocessing Data Text

3.2.1 Seleksi Data

Proses seleksi data merupakan salah satu teknik preprocessing yang sangat penting untuk dilakukan. tahap ini ditujukan untuk mengurangi jumlah atribut yang terkait dalam menentukan target. Dikarenakan atribut data yang berlebihan dan tidak relevan dapat berpengaruh pada proses analisis sentimen dan pada saat proses komputasi, sehingga mempengaruhi hasil nilai akurasi dari proses klasifikasi tersebut (Asmoro et al., 2018).

3.2.2 Cleansing data

Teknik *cleansing* data digunakan untuk membersihkan data dari karakter atau simbol yang tidak relevan atau mengganggu, seperti tanda baca atau emotikon. *Metode cleansing* data ini bertujuan untuk memastikan data yang digunakan dalam analisis sentimen adalah data yang sudah bersih dan siap digunakan.

3.2.3 Lower Case

Penelitian ini juga menggunakan *teknik lower case* untuk merubah semua huruf pada teks menjadi huruf kecil (Pitriyanto et al., 2022). Penggunaan *teknik lower case* pada teks dalam penelitian ini dilakukan untuk merubah semua huruf menjadi huruf kecil guna meningkatkan konsistensi dan kemudahan dalam pemrosesan teks.

3.2.4 Stopwords

Stopword merupakan suatu proses untuk menghapus kata yang dianggap tidak penting seperti contoh kata ‘di’, ‘yang’, ‘dan’, ‘ke’, dan semua kata yang terdapat dalam kamus *stopwords* yang sudah dibuat. Tujuan dari proses ini adalah untuk mengurangi jumlah kata yang disimpan dalam daftar token yang nantinya akan dilakukan proses selanjutnya (Septian et al., 2019).

3.2.5 Tokenized

Tokenizing merupakan sebuah konsep yang mencakup proses sederhana dimana urutan teks dipecah menjadi bagian-bagian yang lebih kecil kemudian dimasukkan sebagai input ke dalam model *Natural language processing* (Azhar et al., 2022). *Tokenizing* adalah proses pemotongan sebuah dokumen menjadi bagian-bagian, yang disebut dengan token. Pada saat bersamaan *tokenizing* juga berfungsi untuk membuang beberapa karakter tertentu yang dianggap sebagai tanda baca (Alita & Isnain, 2020).

3.2.6 Stemming

Tujuan utama dari *stemming* adalah untuk menghilangkan infleksi dari kata-kata sehingga variasi kata yang berbeda dapat diperlakukan sebagai satu kata yang sama selama analisis..*Stemming* sangat bermanfaat dalam pemrosesan bahasa alami seperti pada masalah klasifikasi teks, peringkasan teks (Abidin et al., 2021).

3.2.7 Translate Data

Translate data atau menerjemahkan data adalah proses mengubah data dari satu bentuk atau format ke bentuk atau format lainnya. Proses ini bisa melibatkan berbagai jenis perubahan tergantung pada konteksnya, dalam penelitian ini dengan menggunakan bantuan *library translate* yang dapat dijalankan pada python cukup membantu dalam proses terjemahan data yang akan di analisis sentimen.

3.3 Labeling Data

Labeling data merupakan penerapan pada data mining dengan metode SVM data yang telah didapat harus didefinisikan terlebih dahulu sebagai kalimat yang mengandung nilai positif, netral, dan negatif. Masing-masing data tweet harus diberi label sebagai tweet positif, netral, dan negatif (Syah & Witanti, 2022).

3.5 Visualisasi Data

Visualisasi data merupakan salah satu bagian penting dalam penelitian analisis sentimen. Visualisasi data dapat membantu peneliti untuk lebih mudah memahami dan menginterpretasikan hasil analisis.

4. HASIL DAN PEMBAHASAN

4.1 Hasil Pengumpulan Data

Data awal menampilkan data yang baru diperoleh melalui platform sosial media twitter, data yang diperoleh memiliki data sebanyak 5994 rows dan 15 kolom. Data tersebut diperoleh dengan cara crawling pada platform twitter dengan hashtags #AFCU23 dan #Indonesia_vs_Korea_Selatan Seperti pada gambar 3.

Order ID	Customer Name	Order Date	Product	Quantity	Unit Price	Total Price
10101	John Doe	2023-01-01	Smartphone A	2	500.00	1000.00
10102	Jane Smith	2023-01-02	Smartphone B	1	600.00	600.00
10103	Mike Johnson	2023-01-03	Smartphone C	3	400.00	1200.00
10104	Sarah Davis	2023-01-04	Smartphone D	2	550.00	1100.00
10105	David Wilson	2023-01-05	Smartphone E	1	700.00	700.00
10106	Emily Davis	2023-01-06	Smartphone F	1	800.00	800.00
10107	Chris Evans	2023-01-07	Smartphone G	1	900.00	900.00
10108	Olivia Green	2023-01-08	Smartphone H	1	1000.00	1000.00
10109	Benjamin Blue	2023-01-09	Smartphone I	1	1100.00	1100.00
10110	Ava White	2023-01-10	Smartphone J	1	1200.00	1200.00

Gambar 3. Tampilan Data Awal

4.2 Hasil Preprocessing Data Text

Preprocessing data teks merupakan langkah penting dalam analisis data teks dan pemrosesan bahasa alami (NLP). Tujuan utamanya adalah untuk mengubah data mentah menjadi format yang dapat dianalisis lebih mudah dan efektif. Berikut tahapan preprocessing data teks:

4.2.1 Hasil Seleksi Data

Pada tahap seleksi data menghilangkan kolom yang tidak digunakan dalam penelitian, dengan jumlah kolom awal berjumlah 15 kolom setelah di seleksi menjadi 3 kolom yang akan digunakan. Dapat dilihat pada gambar 4.

Gambar 4. Hasil Seleksi Data

4.2.2 Hasil Cleansing Data & Lowercase

Hasil *cleansing* pada gambar 5 tersebut sudah menampilkan text pada *kolom full_text* yang telah dilakukan penghapusan *mention*, *hastagh*, *url*, *retwt*, *non-alfanumerik*, dan *spasi berlebihan*. Dapat dilihat juga *rows* yang ditampilkan hanya 5647 *rows* hal tersebut disebabkan adanya data yang berbentuk *duplicated* dan *data null* maka dilakukannya cleansing data tersebut. Pada *kolom full_text* telah menunjukkan kalimat atau kata yang telah dilakukan *lowercase* yang membuat kalimat tersebut berada pada tingkatan huruf yang sama seperti (KEMANA) menjadi (kemana).

	full_text	username	created_at
0	gak hanya pas selesai sih selagi pertandingan ...	rzalandi	Fri Apr 26 03:43:07 +0000 2024
1	shin tae yong	cahwonogiri1980	Fri Apr 26 03:42:53 +0000 2024
2	#free11917 ofc we love shin taeyong and kim minj	nabewry	Fri Apr 26 03:42:46 +0000 2024
3	komitmen nyata pak jokowi dalam meningkatkan t...	TugimanTanggulh	Fri Apr 26 03:42:31 +0000 2024
4	orang baik yg sangat dicintai rakyatnya tag sh...	SeperiJokowi	Fri Apr 26 03:42:14 +0000 2024
...			
5989	terimakasih telah mengukir sejarah	AlW444Fady	Thu Apr 25 21:13:17 +0000 2024
5990	lawan korea banyak drama nya	akunnaalter	Thu Apr 25 21:13:16 +0000 2024
5991	selangkah lagi olimpiade mimpi pun belum perna...	Jenukmi01528287	Thu Apr 25 21:13:16 +0000 2024
5992	jangan asal kulum kalo gakenal assalamualaikum...	belompunyabwtl	Thu Apr 25 21:13:12 +0000 2024
5993	alhamdulillah ya allah lolos ke semifinal	narkosut	Thu Apr 25 21:13:11 +0000 2024
5647	rows x 3 columns		

Gambar 5. Hasil Cleansing & Lowercase

4.2.3 Stopwords

Proses *stopwords* dilakukan bertujuan guna memfilter atau menghilangkan kata yang tidak memiliki bobot atau tidak diperlukan dalam penelitian seperti kata di, yang, ke. Pada gambar 6, kata-kata tersebut merupakan kata yang difilter atau dihilangkan dalam penelitian analisis sentiment.

```
[tidak", "bgttttttt", "p", "jg", "jdi", "gabu", "org", "ig", "waaaa", "yang", "pada", "dari", "ke", "di", "itu", "ini", "dgn", "ke", "ga", "jah", "ge", "tgt", "hian", "aku", "kamu", "lu", "list", "gue", "kuk", "laku", "tadi", "ke", "tp", "nih", "yang", "yg", "dah", "tau", "mang", "udh", "dah"]
```

Gambar 6. Hasil Proses stopwords

4.2.4 Tokenized

Tampilan ini merupakan hasil dari *tokenized* yang telah dilakukan pemecahan text menjadi unit-unit yang lebih kecil, *tokenized* ini dapat berupa frasa, kata, dan karakter. Pada penelitian ini *tokenized* di pecah menjadi tiap-tiap kata seperti pada gambar 7.

```
0      [gak, pas, selesai, sih, pertandingan, beberapa...  
1                      [shin, tae, yong]  
2      [free11917, ofc, we, love, shin, taeyong, and,...  
3      [komitmen, nyata, pak, jokowi, meningkatkan, l...  
4      [orang, baik, sangat, dicintai, rakyatnya, tag...  
      ...  
5989      [terimakasih, mengukir, sejarah]  
5990      [lawan, korea, banyak, drama, nya]  
5991      [selangkah, olimpiade, mimpi, belum, pernah, s...  
5992      [jangan, asal, kulum, kalo, gakenal, assalamuala...  
5993      [alhamdulillah, allah, lolos, semifinal]  
Name: full_text, Length: 5647, dtype: object
```

Gambar 7. Hasil Proses Tokenized

4.2.5 Stemming

Kolom full_text dilakukan proses *stemming* dikarenakan *kolom full_text* dijadikan target analisis sentimen, hasil stemming ini merupakan hasil dari proses yang telah dilakukan perubahan kata-kata yang terinfleksi atau terderivasi ke bentuk dasarnya, seperti pada gambar 8.

full_text	
0	gak pas selesai sht banding beberapa teman mu...
1	shin bae yong
2	tree11917 olo we love shin taeyong and kim jong
3	komitmen nyata pak jokowi tingkat tahan senat
4	orang baik sangat cinta rakyat lag shin bae yo...
...	
5642	terimakasih ukir sepanjang
5643	lawan korea banyak drama nya
5644	selang olimpiade mimpi pernah jauh indonesia k
5645	tinggi rasa kultum kalo gakernat assalamualaikum
5646	alhamdulillah allah lohio sentinal
5647	
564 / rows x 1 columns	

Gambar 8. Hasil Proses Stemming

4.2.6 Translate Data

Dilakukannya *translate data* dikarenakan pada proses pengumpulan data, data yang diambil merupakan tanggapan dari seluruh warga net yang turut serta dalam menanggapi laga AFC U23 Indonesia vs Korea selatan. Bahasa yang diperoleh dalam data sentimen memiliki lebih dari 2 bahasa maka dari itu dilakukannya *translate data* ke dalam bahasa inggris dengan menambahkan kolom *tweet_english* sebagai penampung hasil *translate data*. Seperti pada gambar 9.

full_text	tweet_english
gak pas selesai sht banding beberapa teman shin bae yong supporter korea deper ampe senyum suryam	It wasn't until the match was finished that we...
shin bae yong	Shin Bae Yong
tree11917 olo we love shin taeyong and kim jong	tree11917 olo we love shin taeyong and kim jong
komitmen nyata pak jokowi tingkat tahan senat	Pak Jokowi's real commitment to the people of...
orang baik sangat cinta rakyat lag shin bae yo...	good people really love the people lag shin bae...
terimakasih ukir sepanjang	terimakasih ukir sepanjang
lawan korea banyak drama nya	lawan korea banyak drama nya
selang olimpiade mimpi pernah jauh indonesia k	selang olimpiade mimpi pernah jauh indonesia k
tinggi rasa kultum kalo gakernat assalamualaikum	tinggi rasa kultum kalo gakernat assalamualaikum
alhamdulillah allah lohio sentinal	alhamdulillah allah lohio sentinal

Gambar 9. Hasil Translate Data

4.3 Labeling Data

Labeling data dilakukan guna mengetahui hasil dari tiap-tiap tanggapan yang telah diberikan warga net berupa tanggapan *positif*, *negative*, atau *netral*. Pada gambar 10, menampilkan hasil sentimen warganet dengan hasil positif 2587, negatif 738, netral 2322 dan jumlah total tanggapan sebanyak 5647.

Hasil Analisis Data:
Positif = 2587
Netral = 2322
Negatif = 738

Total Data : 5647

Gambar 10. Hasil Labeling Data

4.4 Visualisasi Data

Visualisasi data merupakan proses representasi data dalam bentuk grafis atau gambar, seperti grafik, diagram, atau peta, untuk memudahkan pemahaman dan analisis data. Visualisasi data bertujuan untuk menyajikan data secara visual sehingga pola, tren, dan insight yang tersembunyi dalam data dapat dengan cepat dan mudah dipahami. Berikut hasil visualisasi *wordcloud* dan grafik dari sentimen analisis :

4.4.1 wordcloud

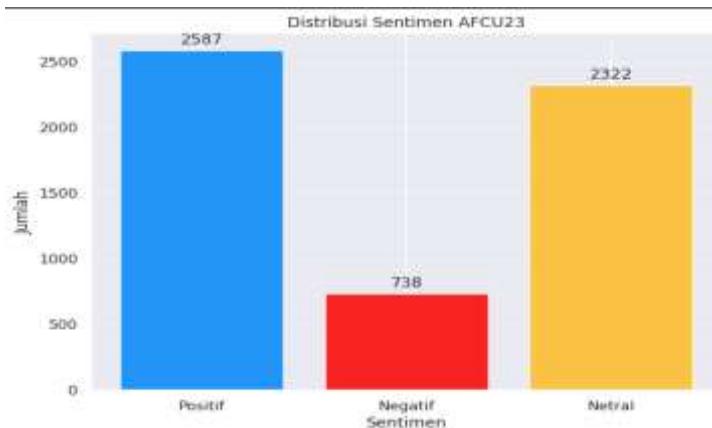
visualisasi *wordcloud* membantu mengidentifikasi kata-kata kunci dengan cepat, dan membuat data teks lebih mudah dipahami, pada hasil *wordcloud* kata Indonesia dan kata *coach sty* pada visualisasi *wordcloud* menampilkan hasil atau tampilan yang besar hal itu di sebabkan mayoritas atau hampir semua warga net sering menyebutkan kata-kata tersebut dalam tanggapan sentiment mereka seperti pada gambar11.



Gambar 11. Hasil Visualisasi Wordcloud

4.4.2 Hasil Diagram Sentiment

Tampilan hasil dari diagram sentimen ini merupakan hasil dari visualisasi dari hasil labeling sentimen sebelumnya, dilakukan visualisasi hasil sentimen dikarenakan dapat lebih mudah dipahami dan dimengerti perbandingan antara sentimen positif, negatif, dan netral. Hasil dari visualisasi ini dapat dilihat pada gambar 12.

**Gambar 12.** Visualisasi Diagram Sentiment

5. KESIMPULAN

Secara keseluruhan, sentimen terhadap laga AFC U23 antara Indonesia dan Korea Selatan cenderung sangat positif. Dengan lebih dari separuh tanggapan bersifat positif dan jumlah sentimen negatif yang cukup rendah, dapat disimpulkan bahwa pertandingan ini diterima dengan baik oleh publik. Sentimen netral yang signifikan juga menunjukkan adanya banyak diskusi obyektif seputar pertandingan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah senantiasa memberikan bimbingan dan saran dalam menyelesaikan penelitian ini. Ucapan terima kasih juga disampaikan kepada STMIK Dharmawacana Metro. kami juga ingin mengucapkan terima kasih kepada Orang Tua yang telah memberikan dukungan dan bantuan dana. Serta tak lupa, terima kasih kepada teman-teman yang telah membantu penulis dalam perjalanan ini.

DAFTAR PUSTAKA

- Abdulloh, F. fauzi, & Pambudi, I. R. (2021). Analisis sentimen pengguna youtube terhadap program vaksin covid-19. ... *Science Research and* <http://csrid.potensi-utama.ac.id/ojs/index.php/CSRID/article/view/524>
- Abidin, Z., Wijaya, A., & Pasha, D. (2021). Aplikasi Stemming Kata Bahasa Lampung Dialek Api Menggunakan Pendekatan Brute-Force dan Pemograman C. *Jurnal Media* <http://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/2483>
- Aggarwal, C. C., & Zhai, C. X. (2012). A survey of text classification algorithms. *Mining Text Data.* https://doi.org/10.1007/978-1-4614-3223-4_6
- Alita, D., & Isnain, A. R. (2020). Pendekripsi Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *Jurnal Komputasi.* <https://jurnal.fmipa.unila.ac.id/komputasi/article/view/2615>
- Ananda, D., & Suryono, R. R. (2024). Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naïve Bayes. *JURNAL MEDIA INFORMATIKA* <http://stmik-budidarma.ac.id/ejurnal/index.php/mib/article/view/7517>

-
- Aryanti, D. (2022). Analisis sentimen ibukota negara baru menggunakan metode naive bayes classifier. *Journal of Information System Research (JOSH)*. <http://ejurnal.seminar-id.com/index.php/josh/article/view/1944>
- Asmoro, A. S. B., Irianto, W. S. G., & ... (2018). Perbandingan Kinerja Hasil Seleksi Fitur pada Prediksi Kinerja Akademik Siswa Berbasis Pohon Keputusan. *JEPIN (Jurnal Edukasi Dan*
<https://jurnal.untan.ac.id/index.php/jepin/article/view/29294>
- Azhar, R., Surahman, A., & Julianne, C. (2022). Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes. *J-SAKTI (Jurnal Sains*
<http://tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/443>
- Azzahra, S. (2021). PERSEPSI PENGGUNA TWITTER TERHADAP KEPRIBADIAN MEREK DAN KORELASINYA DENGAN SOCIAL MEDIA ENGAGEMENT. *Jurnal Gama Societa*, 3(2), 19.
<https://doi.org/10.22146/jgs.63894>
- Hadna, N. M. S., Santosa, P. I., & ... (2016). Studi literatur tentang perbandingan metode untuk proses analisis sentimen di Twitter. In *Semin. Nas. Teknol. Inf* researchgate.net.
https://www.researchgate.net/profile/Nurrun-Muchammad-Hadna/publication/292831965_Studi_Literatur_Tentang_Perbandingan_Metode_Untuk_Proses_Analisis_Sentimen_di_Twitter/links/56b182ec08ae5ec4ed4895b1/Studi-Literatur-Tentang-Perbandingan-Metode-Untuk-Proses-Analisis-Sentimen-di-Twitter.pdf
- Hanafi, M. A., & Solichin, A. (2023). ANALISIS SENTIMEN TERHADAP PSSI ATAS TRAGEDI KANJURUHAN MENGGUNAKAN MULTINOMIAL NAÏVE BAYES. *Telematika MKOM*, 15(1), 21.
<https://doi.org/10.36080/telematikamkom.2292>
- Huang, F., Schintler, L., & McNeely, C. (2019). Data Cleansing. In *Encyclopedia of Big Data*. researchgate.net.
https://www.researchgate.net/profile/Fang-Huang-9/publication/334376431_Data_Cleansing/links/5d2641d292851cf440788b1e/Data-Cleansing.pdf
- Hudha, M., Supriyati, E., & ... (2022). ANALISIS SENTIMEN PENGGUNA YOUTUBE TERHADAP TAYANGAN# MATANAJWAMENANTITERAWAN DENGAN METODE NAÃ VE BAYES CLASSIFIER. *JIKO (Jurnal Informatika* <http://ejurnal.unkhair.ac.id/index.php/jiko/article/view/3376>
- Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Teknika*.
<http://ejurnal.ikado.ac.id/index.php/teknika/article/view/311>

- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*. <https://www.sciencedirect.com/science/article/pii/S0007681309001232>
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information ...* <https://www.nowpublishers.com/article/Details/INR-011>
- Permana, F. C., Wicaksono, Z. M., Kurniawan, C., Abdullah, A. S., & Ruchjana, B. N. (2021). Perception analysis of the Indonesian society on twitter social media on the increase in BPJS kesehatan contribution in the Covid 19 pandemic era. *Journal of Physics: Conference Series*, 1722(1), 012022. <https://doi.org/10.1088/1742-6596/1722/1/012022>
- Pitriyanto, A., Ramadhan, M. N., Al Hakim, T., & Aggry Saputra. (2022). Analisis Sentimen Pembelajaran Daring Pada Masa Pandemi Covid-19 di Indonesia. *Jurnal Bangkit Indonesia*, 11(1), 7–13. <https://doi.org/10.52771/bangkitindonesia.v11i1.197>
- Putra, G. G. S., Swastika, W., & Irawan, P. L. T. (2022). Perbandingan Particle Swarm Optimization dengan Genetic Algorithm dalam Feature Selection untuk Analisis Sentimen pada Permendikbudristek PPKS-LPT. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 8(3), 412. <https://doi.org/10.26418/jp.v8i3.57300>
- Sanderson, J. (2011). *It's a whole new ball game: How social media is changing sports*. Hampton Press.
- Sari, W. P., Mei Ie, M. I., & Tunjungsari, H. K. (2021). Factors Shaping Brand Awareness of Sumba Weaving Products in a Phenomenological Perspective. *Jurnal Komunikasi*, 13(2), 302. <https://doi.org/10.24912/jk.v13i2.11082>
- Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *INSYST: Journal of ...* <https://jurnal.istts.ac.id/index.php/INSYST/article/view/36>
- Syah, H., & Witanti, A. (2022). Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm). *Jurnal Sistem Informasi Dan ...* <http://ejournal.lppm-unbaja.ac.id/index.php/jsii/article/view/1411>
- Tamami, A. B. (n.d.). *THE RIVALRY OF FOOTBALL SUPPORTERS IN INDONESIA AT FANATICISM FRAME OF BONEK AND AREMANIA*.
- Ulfatriyani, H., Nugroho, H. A., & ... (2020). Implementing Term Frequency-Inverse Term Frequency at Tweets in Indonesian Fraud Crime Cases. *2020 3rd International ...* <https://ieeexplore.ieee.org/abstract/document/9331996/>