

ANALISIS SENTIMEN TERHADAP PSSI ATAS TRAGEDI KANJURUHAN MENGGUNAKAN MULTINOMIAL NAÏVE BAYES

Mohammad Afif Hanafi^{1*}, Achmad Solichin²

^{1,2} Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur

Jl. Ciledug Raya, RT.10/RW.2, Petukangan Utara, Kec. Pesanggrahan, Kota Jakarta Selatan, DKI Jakarta 12260

E-mail Koresponden: 1911510038@student.budiluhur.ac.id

(received: 17/2/2023, revised: 12/03/2023, accepted: 14/3/2023)

Abstrak

PSSI (Persatuan Sepakbola Seluruh Indonesia) merupakan suatu federasi sepakbola yang ada di Indonesia. Sebelum dan Sesudah Tragedi Kanjuruhan, PSSI memiliki banyak sekali isu sentimen negatif maupun positif. Oleh karena itu, sentimen ini banyak menimbulkan opini terhadap PSSI sehingga perlu adanya sistem untuk mengetahui setiap kalimat pada sebuah sentimen. Penelitian ini melatarbelakangi opini masyarakat Indonesia baik yang pro ataupun kontra pada Tragedi Kanjuruhan, hal ini menjadi suatu hal yang sangat penting untuk mengetahui sentimen positif atau negatifnya sebuah pendapat yang dilontarkan oleh masyarakat salah satunya melalui *Twitter*. Untuk itu, penelitian ini bertujuan untuk mengklasifikasikan dan menganalisis pandangan dari masyarakat melalui media sosial *Twitter* terhadap PSSI baik sebelum Tragedi Kanjuruhan maupun sesudah Tragedi Kanjuruhan, dan berapa akurasi tertinggi dari algoritme yang diimplementasikan berupa sistem yang berbasis website. Sentimen analisis ini dilakukan melalui proses pelabelan secara manual yang kemudian dibagi menjadi dua kategori yaitu label positif dan label negatif. Dataset bersumber dari *Twitter* dengan proses *scraping* dengan kata kunci sebelum Tragedi Kanjuruhan yaitu "PSSI" pada periode 01 Januari 2022 hingga 31 Agustus 2022 dengan total 452 *tweets*. Dan kata kunci sesudah Tragedi Kanjuruhan yaitu "PSSI", "Sepakbola", "Kanjuruhan", dan "FIFA" pada periode 07 Oktober 2022 hingga 16 Oktober 2022 dengan total 1238 *tweets*. Hasil dari pengujian menggunakan metode *Multinomial Naïve Bayes* mampu menghasilkan analisa pengujian akurasi tertinggi sebelum Tragedi Kanjuruhan bernilai 73% dengan rasio terbaik ialah 80% data latih dan 20% data uji, sedangkan akurasi tertinggi sesudah Tragedi Kanjuruhan bernilai 68% dengan rasio terbaik ialah 90% data latih dan 10% data uji.

Kata Kunci: Kanjuruhan, *Multinomial Naïve Bayes*, Perbandingan Sentimen Analisis, PSSI, Sentimen Analisis, *Twitter*

SENTIMENT ANALYSIS ON PSSI OVER KANJURUHAN TRAGEDY USING MULTINOMIAL NAÏVE BAYES

Abstract

PSSI (Indonesian Football Association) is a football federation in Indonesia. Before and after the Kanjuruhan Tragedy, PSSI had a lot of negative and positive sentiment issues. Therefore, this sentiment raises a lot of opinions about PSSI so that there is a need for a system to find out each sentence in a sentiment. This research is motivated by the opinions of the Indonesian people, both pro and con on the Kanjuruhan Tragedy, this is a very important thing to know the positive or negative sentiment of an opinion expressed by the community, one of which is through *Twitter*. Therefore, this research aims to classify and analyze the views of the public through *Twitter* social media towards PSSI both before and after the Kanjuruhan Tragedy, and what is the highest accuracy of the algorithm implemented in the form of a web-based system. Sentiment analysis is done through a manual labeling process which is then divided into two categories, namely positive labels and negative labels. The dataset is sourced from *Twitter* with a *scraping* process with the keyword before the Kanjuruhan Tragedy, namely "PSSI" in the period January 01, 2022 to August 31, 2022 with a total of 452 *tweets*. And the keywords after the Kanjuruhan Tragedy are "PSSI", "Football", "Kanjuruhan", and "FIFA" in the period October 07, 2022 to October 16, 2022 with a total of 1238 *tweets*. The results of testing using the *Multinomial Naïve Bayes* method were able to produce the highest accuracy test analysis before the Kanjuruhan Tragedy was 73% with the best ratio of 80% training data and 20% test data, while the highest accuracy after the Kanjuruhan Tragedy was 68% with the best ratio of 90% training data and 10% test data.

Keywords: *Comparison of Sentiment Analysis, Kanjuruhan, Multinomial Naïve Bayes, PSSI, Sentiment Analysis, Twitter*

1. Pendahuluan

Analisis sentimen merupakan suatu teknik yang digunakan untuk mengidentifikasi dan mengukur sentimen atau perasaan yang terkandung dalam teks. Dalam dunia maya, analisis sentimen sering digunakan untuk menganalisis perasaan masyarakat terhadap suatu topik atau peristiwa melalui media sosial. Analisis sentimen digunakan untuk menemukan informasi berharga yang dibutuhkan dari data yang tidak terstruktur [1], contohnya ialah *Twitter*. Pada penelitian ini yaitu tentang analisis sentimen terhadap Persatuan Sepakbola Seluruh Indonesia (PSSI) sebelum dan sesudah Tragedi Kanjuruhan.

PSSI merupakan singkatan dari Persatuan Sepakbola Seluruh Indonesia yang merupakan suatu federasi sepakbola yang ada di Indonesia dan khususnya menangani sepakbola Indonesia. Sebelum dan sesudah Tragedi Kanjuruhan, federasi sepakbola Indonesia memiliki banyak sekali sentimen atau pro dan kontra. Oleh karena itu, sentimen ini banyak menimbulkan opini terhadap federasi sepakbola Indonesia sehingga perlu adanya sistem untuk mengetahui setiap kalimat pada sebuah sentimen.

Dengan adanya opini atau pro dan kontra, hal ini menjadi suatu hal yang sangat penting untuk mengetahui sentimen positif atau negatifnya sebuah pendapat yang dilontarkan oleh masyarakat salah satunya melalui media sosial *twitter*. *Twitter* merupakan layanan media sosial berkategori *microblogging* yang paling populer, pengguna juga dapat membaca dan berbagi pesan singkat dengan jumlah maksimal 280 karakter [2]. *Twitter* merupakan media sosial yang menyediakan teks secara singkat sehingga pengguna dapat menuliskan opini atau pendapatnya secara padat, dan jelas dan dapat dipublikasikan ke semua orang atau kelompok tertentu. Kemudian *twitter* menyediakan layanan API untuk pengambilan data [3].

Beberapa penelitian sebelumnya tentang sentimen analisis terhadap PSSI diantaranya penelitian yang pertama adalah membahas tentang Analisis Sentimen Masyarakat Terhadap Kinerja Persatuan Sepakbola Seluruh Indonesia Menggunakan Metode Backpropagation [4], yang menyatakan bahwa melalui pendekatan *development research* meneliti 3 sentimen yaitu kriteria positif dengan prediksi 19,8% sebanyak 198 data, kriteria negatif dengan prediksi 19,5% sebanyak 195 data, dan kriteria netral dengan prediksi 60,7% sebanyak 607 data. Dalam penelitian lain yang berjudul Analisis Sentimen Tentang Opini Performa Klub Sepakbola pada Dokumen *Twitter* menggunakan *Support Vector Machine* dengan Perbaikan Kata Tidak Baku [5], menyatakan bahwa metode *Support Vector Machine* dan *Levenshtein Distance* dengan perbaikan kata tidak baku memiliki akurasi sebesar 83,25%, sedangkan metode *Support Vector Machine* tanpa menggunakan *Levenshtein Distance* perbaikan kata tidak baku memiliki akurasi sebesar 82,75%, yang artinya perbaikan kata tidak baku menggunakan *Levenshtein Distance* terbukti lebih tinggi nilai akurasinya dari pada tanpa menggunakan *Levenshtein Distance*. Dan dalam penelitian lain terhadap sepakbola di Indonesia yaitu dengan judul Analisis Sentimen Pengguna *Twitter* Terhadap Polimik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor [6], menyatakan dengan metode K-Nearest Neighbor dengan nilai $K=23$ memiliki akurasi nilai yaitu 79,99%. Pada penelitian lainnya, Saddam dkk. melakukan analisis persepsi masyarakat terhadap penanganan banjir di DKI Jakarta menggunakan metode *Support Vector Machine* (SVM) [7].

Kontribusi penelitian ini bertujuan untuk mengklasifikasikan opini masyarakat pada sosial media *twitter* terhadap PSSI tentang perbandingan sebelum dan sesudah teragedi kanjuruhan, serta mengetahui nilai akurasi dari metode *Multinomial Naïve Bayes* [8] dengan menggunakan ekstraksi fitur TF-IDF. Dataset yang digunakan diambil dari teks *tweets* yang bersumber dari media sosial *Twitter* dengan kata kunci sebelum Tragedi Kanjuruhan yaitu 'PSSI' pada periode 01 Januari 2022 hingga 31 Agustus 2022 dengan total 452 tweets. Sedangkan kata kunci yang di gunakan sesudah Tragedi Kanjuruhan yaitu 'PSSI', 'Sepakbola', 'Kanjuruhan', dan 'FIFA' pada periode 07 Oktober 2022 hingga 16 Oktober 2022 dengan total 1238 tweets. Hasil analisis tersebut menghasilkan 2 sentimen yaitu Positif dan Negatif.

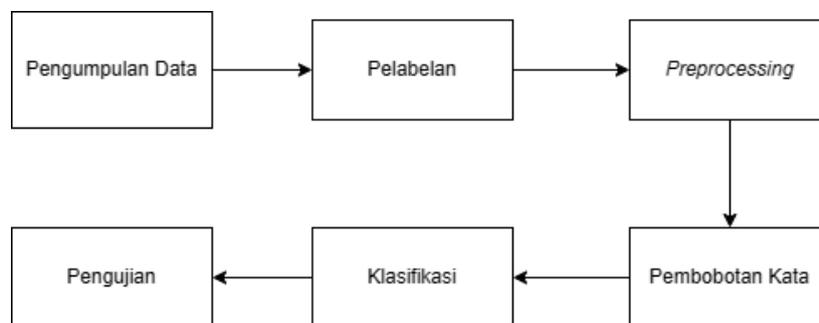
2. Metode Penelitian

2.1. Data Penelitian

Dataset pada penelitian ini yaitu bersumber dari media sosial Twitter. Dataset yang digunakan merupakan perbandingan dari sebelum Tragedi Kanjuruhan dan sesudah Tragedi Kanjuruhan. Dataset tersebut diperoleh menggunakan pustaka *twint git bash* melalui proses *scraping*. Terakumulasi total data sebelum Tragedi Kanjuruhan sebanyak 452 data *tweets* pada periode *tweets* mulai dari 01 Januari sampai 31 Agustus 2022. Data yang diambil menggunakan 1 parameter kata kunci yaitu 'PSSI'. Sedangkan total data sesudah Tragedi Kanjuruhan sebanyak 1238 data *tweets* pada periode *tweets* mulai dari 07 Oktober sampai 16 Oktober 2022. Data yang diambil menggunakan pencarian kata kunci 'PSSI', 'Sepakbola', 'Kanjuruhan' dan 'FIFA'. Lalu data tersebut diolah melalui tahapan *preprocessing* sehingga dapat digunakan sebagai dataset.

2.2. Penerapan Metode

Dalam membangun sistem analisis sentimen yang dilakukan pada penelitian ini menggunakan metode Multinomial Naïve Bayes, terdapat beberapa tahapan yang dilakukan agar penelitian menjadi sesuai dengan tujuan dan terukur. Tahapan tersebut merepresentasikan setiap rancangan dan proses, dari awal hingga akhir sistem berjalan.



Gambar 1. Tahapan Metode

Tahapan yang paling pertama yaitu pengumpulan data dilakukan melalui proses *scraping* untuk mendapatkan data yang bersumber dari media sosial *Twitter*. Pengertian *scraping* adalah mengambil sebuah dokumen semi-terstruktur dari internet dan melakukan analisis pada dokumen tersebut untuk mengekstrak data tertentu dari halaman tersebut, yang nantinya dapat dimanfaatkan untuk kepentingan lain [9]. Penelitian lain juga mengungkapkan bahwa hasil dari proses *scraping* dapat dimanfaatkan oleh sistem lain atau dianalisis lebih lanjut untuk kepentingan lainnya [10].

Data yang diperoleh dalam format CSV, selanjutnya disimpan kedalam basis data (*database*) untuk dilakukan proses *preprocessing* melalui sistem analisis sentimen berbasis web. Pada proses *preprocessing* dilakukan proses penggantian jenis huruf, pembersihan kata *noise*, mengganti dan menghilangkan kata yang tidak layak untuk diproses lebih lanjut dan memberi label kelas sentimen.

Pelabelan sentimen pada penelitian ini, dilakukan secara manual baik sebelum dan juga sesudah Tragedi Kanjuruhan. Pelabelan adalah sebuah proses dimana sebuah dokumen atau kalimat diberikan label kelas berdasarkan ciri-ciri atau karakteristik tertentu. Dalam membagi kelas sentimen, baik sentimen positif maupun negatif, dilakukan dengan pelabelan manual yang melibatkan tiga orang, yaitu seorang pengguna aktif Twitter, seorang penggemar sepak bola, dan seorang mahasiswa. Ketiga orang tersebut melakukan pelabelan dan validasi data bersama-sama. Label positif diberikan kepada *tweet* yang mendukung, menerima, dan percaya pada PSSI. Sementara label negatif diberikan untuk *tweet* yang tidak mendukung, menolak, dan tidak percaya pada PSSI.

Kemudian, pada tahap *preprocessing* yang bertujuan untuk mendapatkan *clean dataset*, sehingga proses pengujian hasil klasifikasi sentimen lebih tepat dan akurat. Adapun proses *preprocessing* yang digunakan pada penelitian ini diantaranya yaitu sebagai berikut:

a. *Case Folding*

Case Folding ialah sebuah proses penyetaraan kata pada suatu huruf besar atau huruf kapital untuk diubah menjadi huruf kecil, misalnya ‘Kanjuruhan’ menjadi ‘kanjuruhan’, ‘PSSI’ menjadi ‘pssi’, dan seterusnya.

b. *Cleansing*

Pada tahapan *cleansing* ini ialah proses pembersihan kata atau menghapus karakter *tweets* yang berisi seperti angka, url, menghapus satu karakter dan simbol lainnya. Seperti contohnya angka 8, URL <https://t.co/KNngkok>, simbol @, menghapus satu karakter contohnya “n”. dan simbol lainnya.

c. Normalisasi

Setelah melalui tahap *cleansing* maka pada penelitian ini dilakukan tahap normalisasi, dimana tahapan ini bertujuan untuk menormalisasikan dari setiap kata yang mengandung kata gaul, kata tidak baku, atau singkatan kata tertentu menjadi kata baku. Pada proses normalisasi ini melibatkan kamus *slangwords* yang terdapat dalam *database*, ditambahkan secara manual dengan kata yang berkaitan dengan sebelum dan sesudah Tragedi Kanjuruhan yang diteliti. Misalnya, kata “dri” diubah menjadi kata “dari”, kata “kmu” menjadi kata “kamu” dan seterusnya.

d. Menghapus *stopwords*

Setelah proses normalisasi, pada penelitian ini yaitu dilanjutkan ke tahap penghapusan *stopwords* atau menghapus kata yang tidak penting dilakukan, biasanya kata umum yang kemunculannya sering ditemui pada dokumen. Proses menghapus *stopwords* ini memanfaatkan daftar dari *stopword* berbahasa Indonesia yang bersumber pada penelitian oleh referensi [11]. Daftar tersebut dilakukan penambahan secara manual dengan kata yang berkaitan dengan sebelum dan sesudah Tragedi Kanjuruhan yang sedang diteliti saat ini. Misalnya, kata “adalah” yang ternyata ada dalam kamus *stopwords* akan dihilangkan dalam suatu dokumen.

e. *Stemming*

Pada tahap berikutnya yaitu tahap *stemming* yang merupakan proses mengubah suatu kata menjadi kata dasar sehingga semua kata imbuhan baik itu awalan, sisipan, akhiran, maupun kombinasi awalan dan akhiran akan dihilangkan. Seperti: kata ‘melihat’ diubah menjadi kata ‘lihat’ dan kata ‘merendahkan’ diubah menjadi kata ‘rendah’. Proses pengubahan tersebut dilakukan dengan melibatkan pustaka Sastrawi yang menerapkan Algoritme Nazief dan Adriani.

3. Hasil dan Pembahasan

3.1 Implementasi Metode

Pada tahap ini merupakan tahapan awal untuk penetapan *dataset*. *Dataset* ini didapatkan melalui proses *scraping* menggunakan pustaka Twint pada periode 01 Januari 2022 hingga 31 Agustus 2022 dengan kata kunci dengan kata kunci sebelum Tragedi kanjuruhan yaitu ‘PSSI’. Kemudian hasil *scraping* disimpan dengan format CSV. *Dataset* ini nantinya akan diimpor ke sistem *database* untuk dilakukan proses lebih lanjut. Contoh data sampel dari total *dataset* sebelum Tragedi Kanjuruhan dapat dilihat pada tabel berikut ini:

Tabel 1. *Dataset* Sebelum Tragedi Kanjuruhan

No.	Created_at	Username	Tweets	Label
1.	2022-08-31 06:58:15	pinem3paska	@Indostransfer Harapan saya Agar Di Sumatera Utara, Kalimantan , dan Irian Jaya, Di tayangkan, beberapa pertandingan ini hanya di Jawa, @vidio @PSSI #Liga2	Negatif
2.	2022-08-31 06:54:57	bhaniie_bsf	@ocheangaluh Haha gimana mau degradasi orang di selametin mulu ama pssi ????	Negatif
3.	2022-08-31 06:53:15	awaveiro	Detail banget woy, sampe pager betis lompat aja dipikiran, nabung woy @PSSI	Negatif

Adapun tahap pengumpulan data sesudah Tragedi Kanjuruhan, *Dataset* ini didapatkan melalui proses *scraping* menggunakan pustaka twint mulai periode 07 Oktober 2022 sampai dengan 16 Oktober

2022 dengan kata kunci sesudah Tragedi kanjuruhan yaitu ‘PSSI’, ‘Sepakbola’, ‘Kanjuruhan’, dan ‘FIFA’. Kemudian hasil *scraping* disimpan dengan format CSV. *Dataset* ini nantinya akan diimpor ke sistem *database* untuk dilakukan proses lebih lanjut. Contoh *dataset* sesudah Tragedi Kanjuruhan ditampilkan pada Tabel 2.

Tabel 2. *Dataset* Sesudah Tragedi Kanjuruhan

No.	Created_at	Username	Tweets	Label
1.	2022-10-16 05:18:19	alhusaini_ asror	@mohmahfudmd Benar dan Betul.Ketua PSSI,mundur secara Sukarela,Sebagai buukti Tanggung Jawab Moral,mundur Secara Sukarela,bukan Secara Hukum.Tegakkan "Hukum dan Moral": GERAKAN MORAL REPUBLIK INDONESIA (GEHUMRI),	Negatif
2.	2022-10-16 05:13:51	Metmalam minggu	@ilhamzada Dilarang intervensi gara" aturan fifa sih , makanya pssi santai" bae , pura" gk denger pssi	Negatif
3.	2022-10-16 05:06:34	onedayas mine	@gibran_tweet Mas.. jadi ketua PSSI yaaa Kalau perlu menpora Sekalian Paragames udah ok apalagi even lain Mosok lobi FIFA pak Menteri BUMN @erickthohir lagi, kelihatan kelasnya soalnya kalau yg lobi menpora atau Iwan bule pasti Indonesia kena sanksi dari FIFA Lobi olimpiade London jg gagal	Negatif

3.2 Preprocessing

Proses tahap *preprocessing* sebelum Tragedi Kanjuruhan tersebut dapat dilihat pada Tabel 3.

Tabel 3. Penerapan *Preprocessing* Sebelum Tragedi Kanjuruhan

Tahap	Teks Awal	Teks proses
<i>Case folding</i>	@SiaranBolaLive Ga habis pikir, PSSI menggunakan sumber daya dari pemerintah menggunakan uang rakyat, menikmati privilege dari pemerintah. Tp pemerintah ga boleh campur tangan. Ada apa sih PSSI dgn FIFA??? Mending pemerintah fokus aja lah sama Bulutangkis	@siaranbolalive ga habis pikir, pssi menggunakan sumber daya dari pemerintah menggunakan uang rakyat, menikmati privilege dari pemerintah. tp pemerintah ga boleh campur tangan. ada apa sih pssi dgn fifa??? mending pemerintah fokus aja lah sama bulutangkis
<i>Cleansing</i>	@siaranbolalive ga habis pikir, pssi menggunakan sumber daya dari pemerintah menggunakan uang rakyat, menikmati privilege dari pemerintah. tp pemerintah ga boleh campur tangan. ada apa sih pssi dgn fifa??? mending pemerintah fokus aja lah sama bulutangkis	ga habis pikir pssi menggunakan sumber daya dari pemerintah menggunakan uang rakyat menikmati privilege dari pemerintah tp pemerintah ga boleh campur tangan ada apa sih pssi dgn fifa mending pemerintah fokus aja lah sama bulutangkis
Normalisasi	ga habis pikir pssi menggunakan sumber daya dari pemerintah menggunakan uang rakyat menikmati privilege dari pemerintah tp pemerintah ga boleh campur tangan ada apa sih pssi dgn fifa mending pemerintah fokus aja lah sama bulutangkis	tidak habis pikir pssi menggunakan sumber daya dari pemerintah menggunakan uang rakyat menikmati hak istimewa dari pemerintah tapi pemerintah tidak boleh campur tangan ada apa sih pssi dengan fifa mending pemerintah fokus aja lah sama bulutangkis
Menghapus Kata Stop	tidak habis pikir pssi menggunakan sumber daya dari pemerintah menggunakan uang rakyat menikmati hak istimewa dari pemerintah tapi pemerintah tidak boleh campur tangan ada apa sih pssi dengan fifa mending pemerintah fokus aja lah sama bulutangkis	habis pikir pssi sumber daya pemerintah uang rakyat menikmati hak istimewa pemerintah pemerintah campur tangan pssi fifa mending pemerintah fokus bulutangkis

Tahap	Teks Awal	Teks proses
<i>Stemming</i>	habis pikir pssi sumber daya pemerintah uang rakyat menikmati hak istimewa pemerintah pemerintah campur tangan pssi fifa mending pemerintah fokus bulutangkis	habis pikir pssi sumber daya perintah uang rakyat nikmat hak istimewa perintah perintah campur tangan pssi fifa perintah fokus bulutangkis

Dan berikut ini proses tahapan dari *preprocessing* sesudah Tragedi Kanjuruhan terdapat pada Tabel 4.

Tabel 4. Penerapan *Preprocessing* Sesudah Tragedi Kanjuruhan

Tahap	Teks Awal	Teks proses
<i>Case folding</i>	kata2 terakhirnya kelas bener "SEKALI LAGI, PSSI BISA APA?" https://t.co/rZqQl0MYok	kata2 terakhirnya kelas bener "sekali lagi, pssi bisa apa?" https://t.co/rzqq10myok
<i>Cleansing</i>	kata2 terakhirnya kelas bener "sekali lagi, pssi bisa apa?" https://t.co/rzqq10myok	kata terakhirnya kelas bener sekali lagi pssi bisa apa
Normalisasi	kata terakhirnya kelas bener sekali lagi pssi bisa apa	kata terakhirnya kelas benar sekali lagi pssi bisa apa
Menghapus Kata <i>Stop</i>	kata terakhirnya kelas benar sekali lagi pssi bisa apa	terakhirnya kelas pssi
<i>Stemming</i>	terakhirnya kelas pssi	akhir kelas pssi

3.1. Pengujian Algoritme

Pada tahap pengujian ini merupakan salah satu hal yang harus dilakukan pada tiap pengembangan proses sebuah sistem untuk melakukan evaluasi tentang tingkat keberhasilan yang telah dilakukan berdasarkan dalam nilai tertentu. Penelitian ini melakukan pengujian baik sebelum Tragedi Kanjuruhan maupun sesudah Tragedi Kanjuruhan berdasarkan perhitungan dari akurasi, presisi dan juga *recall*. Hasil dari pengujian sebelum Tragedi Kanjuruhan terdapat pada Tabel 5.

Tabel 5. Hasil Pengujian Sebelum Tragedi Kanjuruhan

Rasio	90%:10%	80%:20%	70%:30%	60%:40%	50%:50%
Akurasi	70%	73%	65%	66%	64%

Berdasarkan Tabel 5, maka telah diketahui pembagian data dengan rasio 80% data latih dan 20% data uji memiliki nilai akurasi tertinggi yaitu sebesar 73%. Adapun hasil dari pengujian sesudah tragedi Kanjuruhan terdapat pada Tabel 6.

Tabel 6. Hasil Pengujian Sesudah Tragedi Kanjuruhan

Rasio	90%:10%	80%:20%	70%:30%	60%:40%	50%:50%
Akurasi	68%	65%	67%	65%	65%

Berdasarkan Tabel 6, maka telah diketahui pembagian data sesudah tragedi Kanjuruhan dengan rasio 90% data latih dan 10% data uji memiliki nilai akurasi tertinggi yaitu sebesar 68%.

3.2. Perbandingan Sentimen

Pada bagian ini menjelaskan tentang perbandingan dari sentimen masyarakat terhadap PSSI baik sebelum Tragedi Kanjuruhan maupun sesudah Tragedi Kanjuruhan. Pada Tabel 7 adalah hasil akurasi, presisi dan *recall* perbandingan dari sentimen masyarakat terhadap PSSI sebelum dan sesudah Tragedi Kanjuruhan.

Tabel 7. Hasil Akurasi, Presisi, Recall Sebelum dan Sesudah Tragedi Kanjuruhan

	Sebelum Tragedi Kanjuruhan	Sesudah Tragedi Kanjuruhan
Akurasi	73%	68%
Presisi	73%	68%
Recall	73%	69%

Adapun gambar grafik diagram lingkaran dari perbandingan sentimen masyarakat terhadap PSSI sebelum dan sesudah Tragedi Kanjuruhan dibawah ini yaitu dengan keterangan warna hijau berlabel positif dan warna kuning berlabel negatif. Sebelum dan Sesudah Tragedi Kanjuruhan dapat diartikan sentimen negatif lebih banyak dari pada sentimen positif. Gambar dari grafik diagram lingkaran dapat dilihat pada gambar berikut ini:



Gambar 2. Grafik Diagram Lingkaran Perbandingan Sentimen Sebelum dan Sesudah Tragedi Kanjuruhan

4. Kesimpulan

Berdasarkan mengkaji dan menganalisis yang telah dilakukan terhadap sistem sentimen analisis terhadap PSSI sebelum dan sesudah Tragedi Kanjuruhan yang sudah dibuat. Maka dapat diambil kesimpulan untuk membuat sistem sentimen analisis yang lebih baik lagi yaitu: Terdapat dataset sebelum Tragedi Kanjuruhan sebanyak 452 tweets dengan periode 01 Januari 2022 sampai dengan 31 Agustus 2022. Akurasi tertingginya sebesar 73% dengan rasio 80% data latih dan 20% data uji. Sedangkan Terdapat dataset sesudah Tragedi Kanjuruhan sebanyak 1138 tweets dengan periode 07 Oktober 2022 sampai dengan 16 Oktober 2022. Akurasi tertingginya sebesar 68% dengan rasio 90% data latih dan 10% data uji. Sebelum Tragedi Kanjuruhan terdapat sentimen negatif lebih banyak dari pada sentimen positif berdasarkan rasio tertinggi 80% data latih dan 20% data uji, mulai periode 01 Januari 2022 sampai dengan 31 Agustus 2022.

Sedangkan sesudah Tragedi Kanjuruhan terdapat sentimen negatif lebih banyak juga dari pada sentimen positif berdasarkan rasio 90% data latih dan 10% data uji, mulai periode 07 Oktober 2022 sampai dengan 16 Oktober 2022. Pada penerapan proses *case folding*, *cleansing*, normalisasi, menghapus kata stop, dan *stemming* ditahap *preprocessing* sangat berpengaruh dalam meningkatkan akurasi. Sistem sentimen analisis ini dibuat sebagai alat bantu untuk memberikan informasi kepada pihak terkait sebagai bahan pertimbangan dalam menganalisa sentimen dari masyarakat.

Daftar Pustaka

- [1] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *Jurnal Teknoinfo*, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.
- [2] R. Hidayatillah, M. Mirwan, M. Hakam, and A. Nugroho, "Levels of Political Participation Based on Naive Bayes Classifier," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 1, p. 73, 2019, doi: 10.22146/ijccs.42531.

- [3] S. N. J. Fitriyyah, N. Safriadi, and E. E. Pratama, “Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes,” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 3, p. 279, 2019, doi: 10.26418/jp.v5i3.34368.
- [4] R. P. I. Putra, M. Akbar, and R. Amalia, “Analisis Sentimen Masyarakat Terhadap Kinerja Persatuan Sepakbola Seluruh Indonesia Menggunakan Metode Backpropagation,” *Journal of Information Technology Ampera*, vol. 1, no. 2, pp. 106–118, 2020, doi: 10.51519/journalita.volume1.issue2.year2020.page106-118.
- [5] I. Indriati, M. Marji, and S. Pakpahan, “Analisis Sentimen Tentang Opini Performa Klub Sepak Bola Pada Dokumen Twitter Menggunakan Support Vector Machine Dengan Perbaikan Kata Tidak Baku,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 7, pp. 7259–7267, 2019.
- [6] J. A. Septian, T. M. Fahrudin, and A. Nugroho, “Journal of Intelligent Systems and Computation 43,” *Journal of Intelligent Systems and Computation*, pp. 43–49, 2019.
- [7] M. A. Saddam, E. K. Dewantara, and A. Solichin, “Sentiment Analysis of Flood Disaster Management in Jakarta on Twitter Using Support Vector Machines,” *Sinkron*, vol. 8, no. 1, pp. 470–479, Jan. 2023, doi: 10.33395/sinkron.v8i1.12063.
- [8] M. Agus Arianto and A. Solichin, “Analisis Sentimen Motogp Mandalika Pada Twitter Menggunakan Metode Naïve Bayes,” *Jurnal TICOM: Technology of Information and Communication*, vol. 11, no. 1, 2022, [Online]. Available: <https://t.co/XyNW7StiWQ>
- [9] D. D. A. Yani, H. S. Pratiwi, and H. Muhandi, “Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace,” *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, vol. 7, no. 4, p. 257, 2019, doi: 10.26418/justin.v7i4.30930.
- [10] A. Rahmatulloh and R. Gunawan, “Web Scraping with HTML DOM Method for Data Collection of Scientific Articles from Google Scholar,” *Indonesian Journal of Information Systems*, vol. 2, no. 2, pp. 95–104, 2020, doi: 10.24002/ijis.v2i2.3029.
- [11] F. Z. Tala, “A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia,” *M.Sc. Thesis, Appendix D*, vol. pp, pp. 39–46, 2003.