

ANALISIS SENTIMEN OPINI NETIZEN DI INSTAGRAM TERHADAP BERITA SEPAKBOLA LIGA INDONESIA MENGGUNAKAN ALGORITMA LOGISTIC REGRESSION

Fadhil Rifgi Pratama¹, Dahlan²

^{1,2}Politeknik TEDC Bandung alamat; Jl. Politeknik Jl. Pesantren, Cibabat, Kec. Cimahi Utara, Kota Cimahi, Jawa Barat 40513

Keywords:

Analisis Sentimen;
Instagram;
Liga 1 Indonesia;
Logistic Regression;
TF-IDF.

Correspondent Email:

fadhilrifgipratama26@gmail.com

Abstrak. Dengan menggunakan algoritma *Logistic Regression*, penelitian ini bertujuan untuk menganalisis persepsi netizen di Instagram terhadap berita tentang sepak bola Liga Indonesia, khususnya klasemen Liga 1 putaran kedua. Dengan menggunakan teknik *crawling*, data diambil dari komentar pengguna Instagram dan kemudian melalui proses *pre-processing* seperti *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Metode TF-IDF digunakan untuk mengekstraksi karakteristik dan dikategorikan ke dalam tiga kategori sentimen: positif, negatif, serta netral. Hasil penelitian menunjukkan bahwa sebagian besar komentar netral (90,16 persen), dengan skor akurasi, presisi, recall dan skor F1. Model memiliki kemampuan untuk mengklasifikasikan data netral secara dominan dan cukup akurat untuk kategori lain. Hasilnya menunjukkan bahwa *logistic regression* adalah teknik yang efektif untuk mengklasifikasikan opini publik di media sosial tentang masalah olahraga nasional.



JITET is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Abstract. Using the *Logistic Regression* algorithm, this study aims to analyze netizens' perceptions on Instagram towards news about Indonesian League football, especially the second round of League 1 standings. Using the *crawling* technique, data was taken from Instagram user comments and then through *pre-processing* processes such as *cleaning*, *case folding*, *tokenizing*, *stopword removal*, and *stemming*. The TF-IDF method was used to extract characteristics and categorized into three sentiment categories: positive, negative, and neutral. The results showed that most of the comments were neutral (90.16 percent), with accuracy, precision, recall and an F1. The model has the ability to classify neutral data predominantly and is quite accurate for other categories. The results show that *logistic regression* is an effective technique for classifying public opinion on social media about national sports issues.

1. PENDAHULUAN

Sepak bola adalah olahraga yang digemari banyak orang di Indonesia serta menjadi sorotan masyarakat, terutama lewat ajang Liga 1 yang diadakan setiap tahun. Sepanjang musim, pandangan publik mengenai kinerja tim, hasil pertandingan, dan posisi di klasemen sering kali menjadi topik utama di media sosial, termasuk Instagram. Kemajuan teknologi informasi memungkinkan pengguna

internet untuk secara aktif mengungkapkan pendapat dan reaksi terhadap berbagai berita atau konten terkait liga, khususnya pada periode klasemen Liga 1 babak kedua yang mendapat perhatian besar.

Analisis sentimen menjadi salah satu teknik yang signifikan untuk memahami reaksi publik terhadap isu tertentu, termasuk perkembangan dalam dunia sepak bola. Dengan menggunakan data komentar dari Instagram

sebagai sumber utama, penelitian ini bertujuan untuk mempelajari sentimen netizen mengenai berita sepakbola liga indonesia dengan memfokuskan pada klasemen Liga 1 Indonesia putaran kedua dengan menggunakan algoritma *Logistic Regression*. Pendekatan ini dipilih karena kemampuannya untuk secara efisien mengkategorikan data teks ke dalam kelompok sentimen baik itu positif, negatif, ataupun netral, serta dapat diterapkan pada jumlah data yang besar.

Penelitian serupa pernah dilakukan di media sosial Instagram yang berkaitan dengan klub Persija Jakarta selama BRI Liga 1 2022/2023, di mana analisis sentimen berhasil mengungkap pola opini publik dari komentar yang ditulis pada postingan resmi klub [1]. Selain itu, pada penelitian terdahulu lainnya yaitu menganalisis sentimen publik terkait performa tim sepak bola di Indonesia pada aplikasi tiktok [2]. penggunaan algoritma *Logistic Regression* dalam konteks analisis sentimen di Instagram terkait isu klasemen Liga 1, sehingga penelitian ini diharapkan memberikan wawasan baru dalam menganalisis opini publik mengenai perkembangan kompetisi sepak bola nasional.

Diperbandingkan dengan algoritma *k-nearest neighbor* (k-NN), *decision tree*, ataupun model pengklasifikasi lainnya, *logistic regression* memberikan tingkat akurasi yang lebih tinggi. Membandingkan dari algoritma *logistic regression* dengan algoritma yang lainnya juga menunjukkan bahwa algoritma *logistic regression* lebih akurat daripada algoritma k-NN, *support vector machine* (SVM). Studi lain juga memeriksa tiga model pembelajaran mesin *neural network*, *naive Bayes*, dan *logistic regression* untuk mendeteksi diabetes pada 768 data *Kaggle*. Hasil eksperimen menunjukkan bahwa algoritma *logistic regression* memiliki akurasi 75,78%, dibandingkan dengan 74,87% dan 69,27% dari *Naive Bayes* dan *neural network*, masing-masing [3].

Selain itu, algoritma klasifikasi *Logistic Regression* digunakan untuk membuat model hubungan antara variabel independen dengan variabel dependen. Algoritma klasifikasi ini mencari hubungan antara input dengan probabilitas hasil *output*; *input* dapat berupa variabel diskrit atau kontinu, dan probabilitas hasil *output* dari *Logistic*

Regression dapat berupa karakteristik biner (dua kelas) seperti positif atau negatif [4].

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah opini, pemahaman dan teknik analisis yang secara objektif memahami dan menganalisis informasi yang dikemukakan dalam sebuah pernyataan ataupun opini. Metode ini digunakan untuk mengumpulkan, memahami, serta menganalisis data tekstual untuk memastikan perasaan baik itu positif, negatif, atau pendapat netral [5]. Analisis sentimen adalah penelitian pemrosesan data terhadap opini, sentimen, dan emosi yang diungkapkan melalui teks. Tugas utama analisis sentimen adalah mengidentifikasi polaritas dalam teks yang ada dalam dokumen, kalimat, atau pernyataan [6].

2.2 Media Sosial Instagram

Instagram adalah aplikasi berbagi postingan foto yang memberikan pengguna kemampuan untuk mengambil foto, menerapkan filter digital, lalu membagikan foto ini pada berbagai layanan jejaring sosial, termasuk Instagram itu sendiri. Blogging foto dalam jejak orisinal Youtube dan sebuah halaman seperti buku harian online, kemudian juga bisa berlangganan akun pengguna lain. Ini adalah sistem sosial dari instagram pengguna disebut pengguna lain atau pengikut satu sama lain. Pengguna Instagram juga dapat menyukai dan mengomentari foto yang di unggah pengguna dan foto yang diunggah oleh orang lain. Juga banyak dari penggerak dari jumlah suka diunggah oleh pengguna, apakah foto tersebut dikenal atau tidak. Untuk mendapati teman yang menggunakan Instagram, dapat menggunakan teman-teman mereka juga yang menggunakan sosial media seperti Facebook dan Twitter [7].

2.3 Data Mining

Suyanto mendefinisikan data mining sebagai proses penemuan pola-pola yang baru dari kumpulan data yang sangat besar. Ini termasuk metode yang merupakan kombinasi dari *artificial intelligence*, *machine learning*, *statistics*, dan *database systems*. Berdasarkan

definisi para ahli yang disampaikan di atas, maka dapat disimpulkan bahwa data mining adalah proses pencarian data secara otomatis yang memiliki kemampuan untuk mengidentifikasi pola-pola baru dalam kumpulan data yang sangat besar [8]. Data mining adalah langkah untuk menemukan pola ataupun informasi yang menarik dalam sejumlah data yang dipilih dengan menggunakan metode atau teknik tertentu [9].

2.4 Algoritma Logistic Regression

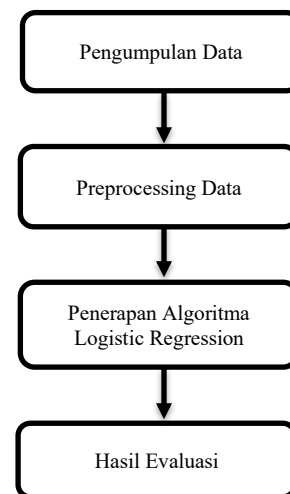
Algoritma *logistic regression* adalah konsep yang sangat penting. *Logistic Regression*, berdasarkan variabel yang relevan, digunakan dalam analisis data untuk memprediksi kemungkinan terjadinya suatu peristiwa atau kejadian. Bentuk fungsi daripada algoritma ini mirip dengan regresi linear, tetapi itu digunakan untuk masalah klasifikasi yang menghasilkan keluaran dalam bentuk probabilitas atau kelas diskret. Kata "binomial" berarti hanya ada dua hasil yang mungkin. Oleh karena itu, *logistic regression* dapat digunakan untuk menyelesaikan berbagai masalah klasifikasi [10]. *Logistic regression* adalah jenis regresi analisis yang menjelaskan hubungan antara variabel dependen dan variabel independen. Ini menjelaskan hubungan antara satu ataupun lebih variabel bebas dan variabel terikat dari jenis kategori tertentu, yang dapat berkisar antara 0 dan 1 dan benar atau salah, besar atau kecil. Jenis variabel bebasnya adalah kategori. Ini adalah perbedaan utama antara regresi logistik dan regresi berganda atau regresi linear lainnya [11].

2.5 Klasifikasi

Menurut Kamus Besar Bahasa Indonesia (KBBI), klasifikasi berarti penyusunan bersistem di dalam kelompok ataupun golongan menurut kaidah ataupun standar yang ditetapkan. Namun, secara umum, klasifikasi adalah proses mengelompokkan benda-benda yang memiliki beberapa karakteristik yang sama dan juga memisahkan benda-benda yang tidak sama [12].

3. METODE PENELITIAN

Tahapan dalam mengolah data di dalam algoritma *logistic regression* terdiri dari beberapa tahap yaitu terdapat pada gambar 3.1.



Gambar 3.1 Bagan Alir Penelitian

3.1 Crawling Data

Data penelitian ini dikumpulkan dari postingan klasemen instagram bri liga 1 menggunakan alat ekspor komentar IG *Chrome Extension* dari 17 Februari 2025 hingga 25 Mei 2025 totalnya adalah 15676 data.

Tabel 1 Hasil *Crawling*

No	Data Set
1	Gabisa ngejar poin,bisanya ngurangin pemain 🤔🤔🤔🤔🤔🤔
2	semakin ku kejar semakin kau jauhhhh
3	percepat sib raih juaranya klo bisa sebelum lebaran

3.2 Pre- Processing

Pre-processing adalah tahapan dalam membersihkan dan menyiapkan data untuk diproses, termasuk menghilangkan kata-kata yang tidak perlu, lalu memecah teks menjadi kata-kata, dan merubah kata ke bentuk dasar.

Tabel 2 Hasil *Cleaning*

Sebelum	Sesudah
Sepertinya klo 20 tim lebih ideal	Sepertinya klo tim lebih ideal

Tabel 3 Hasil *Case Folding*

Sebelum	Sesudah
Bismilah persib juara	bismilah persib juara

Tabel 4 *Tokenizing*

Sebelum	Sesudah
Wah parah nih belum ada kabar laga kemarin, gak ada sanksi juga ya parah	['Wah', 'parah', 'nih', 'belum', 'ada', 'kabar', 'laga', 'kemarin', 'gak', 'ada', 'sanksi', 'juga', 'ya', 'parah']

Tabel 5 *Stopword*

Sebelum	Sesudah
Wah parah nih belum ada kabar laga kemarin gak ada sanksi juga ya parah	Wah parah nih kabar laga kemarin gak sanksi parah

Tabel 6 *Stemming*

Sebelum	Sesudah
Akankah Persib mempertahankan tahta juaranya	akan persib tahan tahta juara

3.3 Penerapan TF-IDF

Tf-IDF adalah teknik pra-pemrosesan teks yang penting yang membantu mengubah kata-kata menjadi angka berdasarkan tingkat kepentingannya. Teknik ini memungkinkan algoritma seperti *logistic regression* melakukan klasifikasi teks secara akurat dan efisien.

3.4 Confusion Matrix

Confusion matrix adalah sebuah tabel yang diterapkan untuk mengevaluasi kinerja pada model klasifikasi *machine learning*. Pada Tabel ini yaitu membandingkan dari hasil prediksi model dengan label sebenarnya (*ground truth*) dan dapat membantu kita untuk memahami kesalahan apa yang dilakukan oleh model.

3.5 Evaluasi

Proses menilai kinerja model pembelajaran mesin terhadap data uji dikenal sebagai evaluasi pada data. Ini dilakukan dalam konteks klasifikasi dengan menghitung metrik seperti "akurasi, presisi, *recall*, dan *confusion*

matrix", yang menunjukkan seberapa baik prediksi model sesuai dengan data sebenarnya. Sangat penting untuk melakukan evaluasi ini untuk memastikan bahwa model sesuai dengan data latih dan bekerja dengan baik pada data baru yang memang belum pernah dilihat sebelumnya.

4. HASIL DAN PEMBAHASAN

4.1 Hasil *Crawling*

Data yang telah dihimpunkan sebanyak 15676 data pada postingan instagram klasemen bri liga 1 putaran ke 2 dari waktu atau periode tertentu. Pengambilan data ditunjukkan pada gambar berikut.

Index	Task	Type	Extracted	
1	DirEzLiyN6H	Comments	907	csv excel
2	DI_rm5pybme	Comments	1815	csv excel
3	DJRiegeyRvD	Comments	2391	csv excel
4	DJilz70SE7s	Comments	1125	csv excel
5	DJ0foK8y4M6	Comments	914	csv excel

Gambar 4.1 Hasil *Crawling*

4.2 Tahapan *Pre-Processing*

Preprocessing adalah pemrosesan data untuk *cleaning*, *case folding*, *tokenizing*, *stopword*, dan *stemming* pada data yang sudah diambil dengan cara *crawling*.

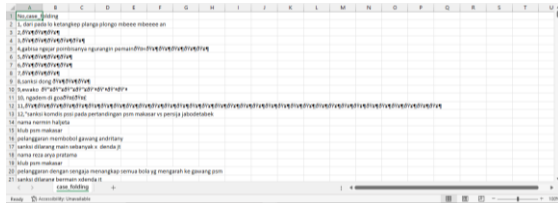
4.2.1 *Cleaning*

Cleaning adalah proses membersihkan elemen teks yang tidak diperlukan yaitu seperti adanya simbol, angka, emoji, atau tanda baca adalah apa yang dimaksudkan untuk membuat data lebih rapi dan siap untuk analisis.

Gambar 4.2 Hasil *Cleaning*

4.2.2 Case Folding

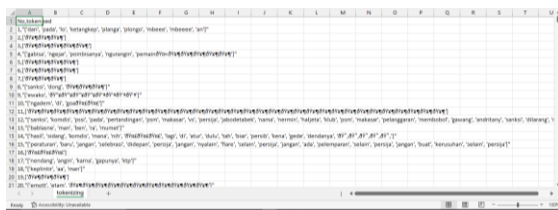
Case Folding adalah cara merubah kalimat dari huruf kapital untuk menjadikannya huruf kecil, Karena arti kata dengan huruf kapital sama dengan kata dengan huruf kecil [13].



Gambar 4.3 Hasil Case Folding

4.2.3 Tokenizing

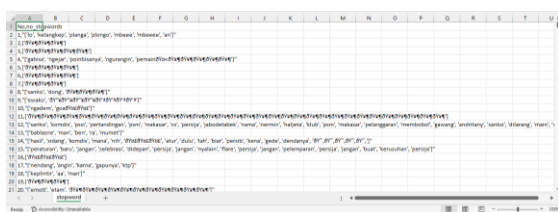
Tokenizing adalah proses membagi teks yang menjadi bagian lebih kecil yang disebut token, yang biasanya yaitu terdiri dari kata-kata individu [14].



Gambar 4.4 Hasil Tokenizing

4.2.4 Stopword

Stopword adalah kata-kata umum yang seringkali muncul dalam teks tetapi dianggap tidak perlu dalam analisis. Dalam pemrosesan bahasa alami, *stopword* biasanya dihapus untuk memfokuskan analisis pada kata-kata yang lebih relevan dan bermakna.

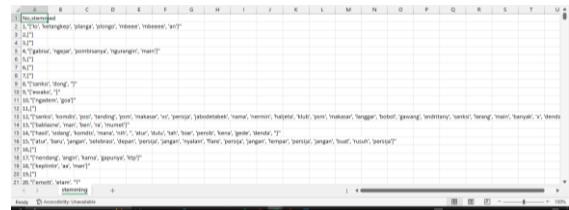


Gambar 4.5 Hasil Stopword

4.2.5 Stemming

Stemming adalah proses untuk merubah kata menjadi kata dasar dengan menghilangkan imbuhan. Sehingga memungkinkan untuk

mengurangi ukuran kata yang berbeda bentuk tetapi tetap memiliki makna yang sama [15].

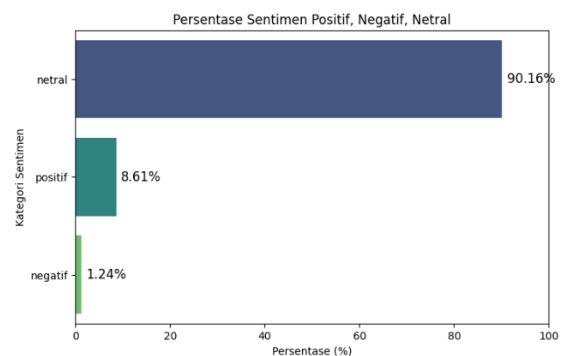


Gambar 4.6 Hasil Stemming

4.6 Hasil Evaluasi

4.6.1 Hasil Grafik Pelabelan Data

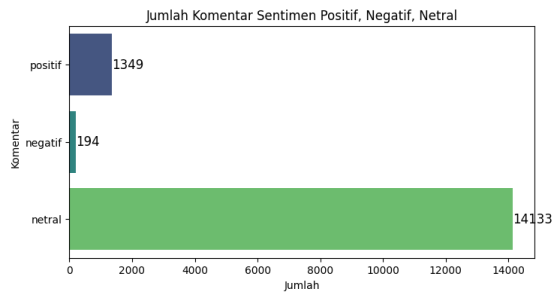
Gambar di bawah ini menunjukkan distribusi persentase sentimen netizen terhadap berita sepakbola Liga Indonesia yang dianalisis menggunakan algoritma *Logistic Regression*. Hasilnya menunjukkan bahwa komentar dengan sentimen netral mendominasi sebesar 90,16%, menunjukkan bahwa sebagian besar pengguna Instagram memberikan komentar yang bersifat informasi daripada opini emosional, seperti pujian atau kritik.



Gambar 4.7 Persentase Sentimen

4.6.2 Jumlah Data Positif, Negatif dan Netral

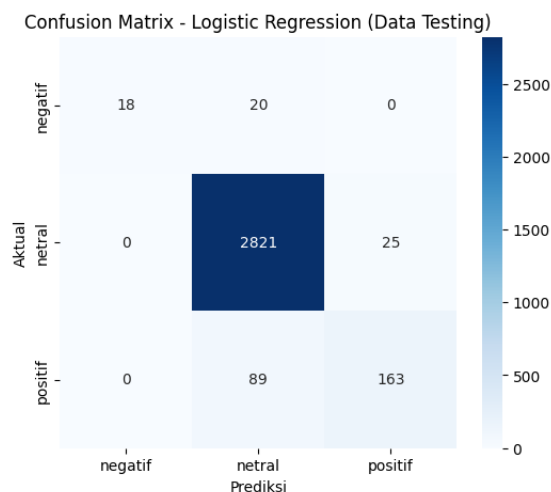
Gambar di bawah ini menunjukkan jumlah komentar netizen yang diposting di Instagram terkait berita sepakbola Liga Indonesia berdasarkan kategori sentimen netral, positif, dan negatif. Jumlah komentar positif berjumlah 1.349, dan yang negatif hanya 194.



Gambar 4.8 Jumlah Data Positif, Negatif dan Netral

4.6.3 Confusion Matrix Logistic Regression

Gambar *confusion matrix* di bawah ini menunjukkan bahwa performa model *Logistic Regression* dalam mengklasifikasikan sentimen komentar netizen di aplikasi Instagram terhadap berita sepakbola Liga Indonesia pada data pengujian. Model mampu mengklasifikasikan komentar netral dengan cukup baik.

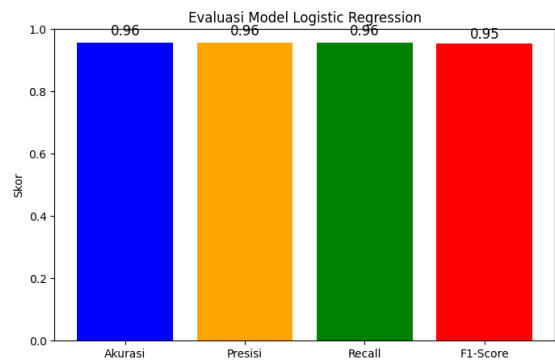


Gambar 4.9 Confusion Matrix

4.6.4 Hasil Evaluasi Model Logistic Regression

Hasil dari evaluasi pada model *logistic regression* ditunjukkan pada gambar di bawah ini skor akurasi, presisi, serta recall yang diperoleh sebesar 0,96, dan skor F1 sebesar 0,95. Nilai-nilai ini menunjukkan bahwa model dapat secara efektif dan konsisten mengklasifikasikan perasaan netizen tentang berita tentang sepak bola Liga Indonesia. Jumlah skor evaluasi yang tinggi menunjukkan bahwa metodologi yang digunakan dalam penelitian mulai dari tahap preprocessing hingga pemilihan algoritma valid. Meskipun

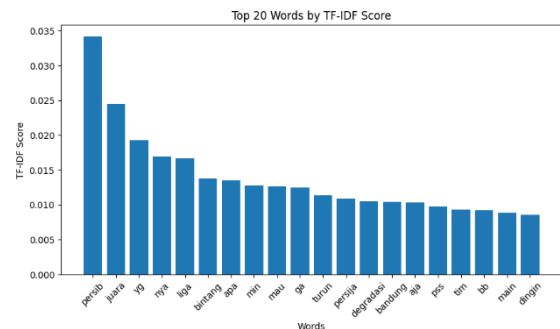
sentimen netral mendominasi data, model masih dapat mengidentifikasi dan memprediksi sentimen pada tingkat akurasi yang tinggi. Maka dengan temuan ini menunjukkan bahwa *Logistic Regression* dapat diandalkan sebagai teknik untuk mengklasifikasikan sentimen dalam konteks media sosial.



Gambar 4.10 Hasil Evaluasi Model Logistic Regression

4.6.5 Top Words TF-IDF

Gambar dibawah ini menampilkan 20 kata teratas dari hasil skor TF-IDF yang merepresentasikan kata-kata paling signifikan pada komentar netizen terhadap berita sepakbola Liga 1 Indonesia.

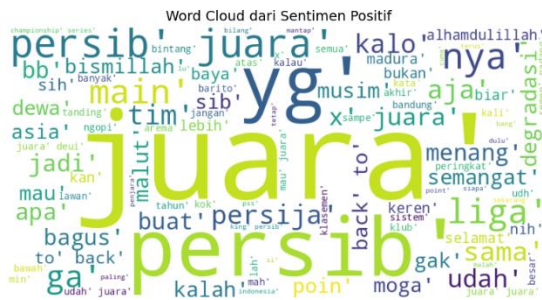


Gambar 4.11 Top Words TF-IDF

4.6.6 Word Cloud TF-IDF Positif

Gambar di bawah ini menunjukkan *word cloud* kata-kata komentar berlabel sentimen positif, yang menunjukkan bahwa kata-kata yang seringkali digunakan oleh netizen untuk memberikan ulasan positif tentang berita sepak bola Liga Indonesia. Terbanyak digunakan adalah kata "juara" dan "persib", yang menunjukkan dukungan kuat dan harapan besar terhadap klub Persib Bandung. Kata-kata seperti "main", "menang", "bagus", dan

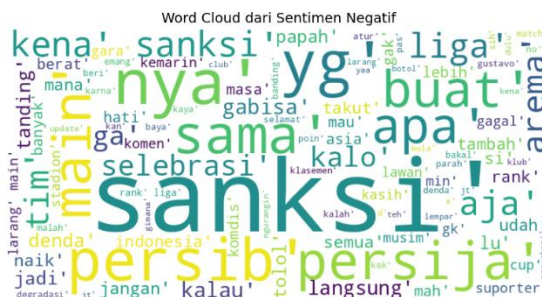
"semangat" menunjukkan semangat, optimisme, dan kegembiraan netizen. *Word cloud* ini mendukung hasil analisis bahwa komentar positif banyak diisi dengan dukungan dan harapan untuk performa tim favorit. Ini juga menunjukkan bahwa model berhasil menangkap sentimen positif dengan benar.



Gambar 4.12 Word Cloud TF-IDF Positif

4.6.7 Word Cloud TF-IDF Negatif

Gambar di bawah ini menunjukkan *word cloud* kata-kata negatif dari komentar netizen, yang menampilkan kata-kata yang paling sering banyak dipakai untuk kritik atau ketidakpuasan mereka terhadap berita tentang sepak bola Liga Indonesia. Untuk menunjukkan bahwa masalah hukuman atau penalti menjadi fokus utama komentar negatif, kata "sanksi" muncul paling sering. Selain itu, kata-kata seperti "persib", "persija", "kena", "main", dan "denda" juga sering muncul, menunjukkan ketidaksetujuan terhadap keputusan tim atau penyelenggara liga. Ketika terjadi kontroversi, pelanggaran, atau ketidakadilan dalam pertandingan, komentar negatif cenderung muncul, dan ini memperkuat hasil klasifikasi sentimen yang dapat secara akurat menangkap nuansa kritik publik.



Gambar 4.13 Word Cloud TF-IDF Negatif

4.6.8 Word Cloud TF-IDF Netral

Gambar di bawah ini merupakan word cloud dari komentar dengan sentimen netral berdasarkan skor TF-IDF sentimen netral, yang menampilkan kata-kata yang paling sering muncul tanpa kecenderungan emosional positif atau negatif. Kata “persib”, “liga”, dan “nya” mendominasi, menandakan bahwa komentar netral banyak berisi informasi atau pernyataan umum terkait tim dan kompetisi. Kata-kata seperti “turun”, “klasemen”, “degradasi”, dan “bintang” juga muncul cukup sering, menggambarkan pembahasan seputar performa tim, posisi di klasemen, atau kondisi pertandingan tanpa ekspresi opini yang kuat. Word cloud ini menguatkan bahwa komentar netral bersifat deskriptif dan faktual, serta menunjukkan bahwa model berhasil mengidentifikasi konten yang tidak menunjukkan sentimen jelas secara akurat.



Gambar 4.14 Word Cloud TF-IDF Netral

5. KESIMPULAN

Studi ini menunjukkan bahwa algoritma *Logistic Regression* berfungsi dengan baik untuk mengklasifikasikan sentimen netizen terhadap berita Liga 1 di Instagram. Model dapat secara akurat membedakan sentimen positif, negatif, dan netral dengan preprocessing dan penerapan TF-IDF. Hasil ini menunjukkan bahwa *logistic regression* dapat menjadi algoritma yang efektif untuk menganalisa komentar-komentar publik netizen di sosial media, terutama mengenai hal sepak bola nasional.

DAFTAR PUSTAKA

- [1] A. M. Jon, "Analisis Sentimen Pada Media Sosial Instagram Klub Persija Jakarta Menggunakan Metode Naive Bayes,"

- Automata*, vol. 958, pp. 1–8, 2022.
- [2] I. Apryani, A. Fauzi, D. S. Kusumaningrum, and H. H. Handayani, “Analisis Sentimen Performa Timnas Sepak Bola Indonesia pada Kolom Komentar Aplikasi TikTok Menggunakan Algoritma Machine Learning,” vol. 8, no. 1, pp. 76–89, 2025.
 - [3] Erlin, Yulvia Nora Marlim, Junadhi, Laili Suryati, and Nova Agustina, “Deteksi Dini Penyakit Diabetes Menggunakan Machine Learning dengan Algoritma Logistic Regression,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 11, no. 2, pp. 88–96, 2022, doi: 10.22146/jnteti.v11i2.3586.
 - [4] B. Ramadhani and R. R. Suryono, “Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse,” *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 714, 2024, doi: 10.30865/mib.v8i2.7458.
 - [5] M. G. Andriawan and T. Ernawati, “Penggunaan Algoritma Naïve Bayes Dan Support Vector Machine Untuk Analisis Sentimen Konflik Palestina Dan Israel Pada Platform X,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, pp. 3222–3230, 2024, doi: 10.23960/jitet.v12i3.4943.
 - [6] T. Juniardi and C. A. Sugianto, “ANALISIS SENTIMEN TIM NASIONAL SEPAK BOLA INDONESIA DI TURNAMEN PIALA DUNIA U-17 INDONESIA PADA TWITTER (X) MENGGUNAKAN,” vol. 12, no. 3, 2024.
 - [7] I. Irwandani and S. Juariyah, “Pengembangan Media Pembelajaran Berupa Komik Fisika Berbantuan Sosial Media Instagram sebagai Alternatif Pembelajaran,” *J. Ilm. Pendidik. Fis. Al-Biruni*, vol. 5, no. 1, pp. 33–42, 2016, doi: 10.24042/jpifalbiruni.v5i1.103.
 - [8] S. P. Dewi, N. Nurwati, and E. Rahayu, “Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor,” *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
 - [9] L. Suryadi, N. Ngajiyanto, N. E. Pratiwi, F. Ardhy, and P. Riswanto, “Penerapan Data Mining Prediksi Penjualan Mebel Terlaris Menggunakan Metode K-Nearest Neighbor(K-Nn) (Studi Kasus : Toko Zerita Meubel),” *JUSIM (Jurnal Sist. Inf. Musirawas)*, vol. 7, no. 2, pp. 174–184, 2022, doi: 10.32767/jusim.v7i2.1697.
 - [10] Y. A. D. K. A’yunan, U. Indahyanti, and S. Busono, “Implementasi Data Mining dalam Klasifikasi Diagnosa Kanker Payudara menggunakan Algoritma Logistic Regression,” *J. TEKINKOM*, vol. 6, no. 2, pp. 400–407, 2023, doi: 10.37600/tekinkom.v6i2.948.
 - [11] F. D. Pramakrisna, F. D. Adhinata, and N. A. F. Tanjung, “Aplikasi Klasifikasi SMS Berbasis Web Menggunakan Algoritma Logistic Regression,” *Teknika*, vol. 11, no. 2, pp. 90–97, 2022, doi: 10.34148/teknika.v11i2.466.
 - [12] J. S. Hutagalung and Rasiban, “Analisis Sentimen Keuangan (Data Fiqa and Financial Phrasebank) Menggunakan Algoritma Logistic Regression Dan Support Vector Machine,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 3, pp. 1654–1669, 2023, doi: 10.35870/jimik.v4i3.404.
 - [13] A. Nayla, C. Setianingsih, and B. Dirgantoro, “Deteksi Hate Speech Pada Twitter,” *eProceeding Eng.*, vol. 10, no. 1, p. 256, 2023.
 - [14] A. Amelia and R. Yusuf, “ANALISIS SENTIMEN MASYARAKAT INDONESIA PADA PLATFROM X TERHADAP ISU FUFUFAFA MENGGUNAKAN BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS,” pp. 72–80, 2025.
 - [15] S. Shevira, I. M. A. D. Suarjaya, and P. W. Buana, “Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia,” *JITTER J. Ilm. Teknol. dan Komput.*, vol. 3, no. 2, p. 1074, 2022, doi: 10.24843/jtrti.2022.v03.i02.p06.