Analisis Sentimen Instagram Persija Jakarta Menggunakan Metode Support Vector Machine (SVM)

Yoseph Yandi Bria[1*], Aprilisa Arum Sari[2], Ridwan Dwi Irawan[3]

[1]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[1*]220103214@mhs.udb.ac.id

[2] Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[2]aprilisa_arumsari@udb.ac.id

[3] Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
[3]ridwan_dwiirawan@udb.ac.id

Abstrak— **Media sosial Instagram telah menjadi platform utama bagi pendukung klub sepak bola dalam mengekspresikan opini dan emosi mereka. Tingginya volume komentar pada akun Instagram Persija Jakarta membutuhkan pendekatan otomatis untuk memahami sentimen publik secara efisien. Penelitian ini bertujuan untuk mengetahui tingkat akurasi metode Support Vector Machine (SVM) dalam mengklasifikasikan sentimen komentar pada akun Instagram resmi Persija Jakarta ke dalam tiga kategori, yaitu positif, negatif, dan netral. Data yang digunakan terdiri dari 6.000 komentar yang dikumpulkan dari 60 postingan, yaitu 30 postingan saat Persija menang dan 30 postingan saat kalah, dalam rentang waktu 1 Juli 2023 hingga 23 Mei 2026. Data kemudian diproses melalui tahapan preprocessing yang meliputi case folding, cleaning, normalisasi, tokenisasi, stopword removal, dan stemming, serta pelabelan sentimen secara otomatis menggunakan kamus INSET Lexicon yang kemudian divalidasi secara manual. Setelah proses tersebut, diperoleh 4.477 komentar yang digunakan dalam penelitian. Ekstraksi fitur dilakukan menggunakan metode TF-IDF yang menghasilkan 4.440 fitur. Dataset dibagi menjadi 3.581 data latih dan 896 data uji, kemudian diklasifikasikan menggunakan SVM dengan kernel linear. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 93,64%, dengan precision tertinggi sebesar 98% pada kelas positif dan recall tertinggi sebesar 99% pada kelas netral. Penelitian ini membuktikan bahwa metode SVM dengan kernel linear mampu mengklasifikasikan sentimen komentar Instagram berbahasa Indonesia secara efektif dengan tingkat akurasi yang tinggi pada konteks komunitas sepak bola Persija Jakarta.**
Kata kunci— Analisis Sentimen, Support Vector Machine, SVM, Instagram, Persija Jakarta, TF-IDF, Natural Language Processing.

Abstract— **Instagram social media has become the primary platform for football club supporters to express their opinions and emotions. The high volume of comments on Persija Jakarta's Instagram account requires an automated approach to efficiently understand public sentiment. This study aims to determine the accuracy level of the Support Vector Machine (SVM) method in classifying comment sentiments on Persija Jakarta's official Instagram account into three categories: positive, negative, and neutral. The data used consists of 6,000 comments collected from 60 posts, comprising 30 posts when Persija won and 30 posts when they lost, within the period of July 1, 2023, to May 23, 2026. The data was then processed through preprocessing stages including case folding, cleaning, normalization, tokenization, stopword removal, and stemming, as well as automated sentiment labeling using the INSET Lexicon dictionary, which was subsequently validated manually. After this process, 4,477 comments were obtained to be used in the study. Feature extraction was performed using the TF-IDF method, generating 4,440 features. The dataset was divided into 3,581 training data and 896 testing data, then classified using SVM with a linear kernel. The evaluation results show that the model achieves an accuracy of 93.64%, with the highest precision of 98% in the positive class and the highest recall of 99% in the neutral class. This study proves that the SVM method with a linear kernel is capable of effectively classifying Indonesian-language Instagram comment sentiments with a high level of accuracy within the context of the Persija Jakarta football community.**
Keywords— Sentiment Analysis, Support Vector Machine, SVM, Instagram, Persija Jakarta, TF-IDF, Natural Language Processing.

I. PENDAHULUAN

Instagram merupakan salah satu platform media sosial dengan pertumbuhan pengguna yang signifikan di Indonesia. Berdasarkan data GoodStats [1], jumlah pengguna Instagram di Indonesia terus mengalami peningkatan dari tahun ke tahun sejak 2018 hingga 2025, menjadikan Indonesia sebagai salah satu negara dengan basis pengguna Instagram terbesar di dunia. Besarnya jumlah pengguna ini mendorong Instagram menjadi ruang publik digital yang aktif digunakan masyarakat untuk berbagi informasi, pendapat, hingga ekspresi emosional terhadap berbagai topik, termasuk olahraga.

Persija Jakarta merupakan salah satu klub sepak bola profesional di Indonesia yang memiliki jumlah pengikut Instagram yang besar dan tingkat interaksi pengguna yang tinggi. Para pendukung kerap memberikan komentar pada setiap unggahan di akun resmi Persija Jakarta, baik saat tim meraih kemenangan maupun kekalahan. Komentar-komentar tersebut mencerminkan berbagai bentuk sentimen, kekecewaandengan karakteristik linguistik yang khas: penggunaan bahasa informal,

slang komunitas, dan ekspresi emosional yang berfluktuasi tajam bergantung pada hasil pertandingan. Fenomena ini menjadikan kolom komentar Instagram Persija Jakarta sebagai sumber data yang kaya untuk memahami opini publik terhadap performa tim [2][3].

Tingginya volume komentar yang masuk setiap harinya menjadi tantangan tersendiri apabila analisis dilakukan secara manual. Dibutuhkan waktu, tenaga, dan sumber daya yang besar untuk membaca dan mengkategorikan ribuan komentar secara satu per satu. Oleh karena itu, diperlukan pendekatan otomatis berbasis kecerdasan buatan yang mampu menganalisis dan mengklasifikasikan sentimen dari teks secara cepat dan efisien [4]. Analisis sentimen merupakan cabang dari Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi dan mengekstrak opini atau perasaan yang terkandung dalam sebuah teks, umumnya diklasifikasikan ke dalam kategori positif, negatif, dan netral [5][6]. Meskipun telah banyak diterapkan di berbagai domain, kajian analisis sentimen pada komunitas suporter sepak bola Indonesia masih sangat terbatas, padahal hasilnya bernilai strategis bagi manajemen klub dan pengelola kompetisi.

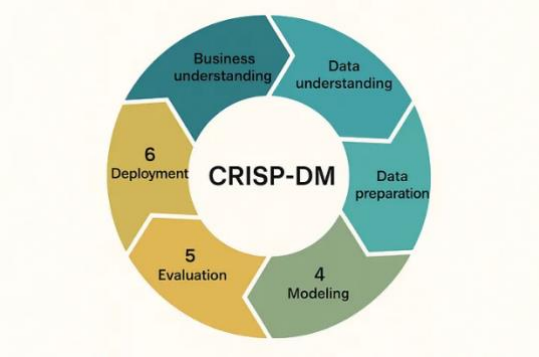
Berbagai metode telah digunakan dalam penelitian analisis sentimen, salah satunya adalah Support Vector Machine (SVM). SVM merupakan algoritma machine learning yang bekerja dengan mencari hyperplane optimal untuk memisahkan kelas data, dan telah terbukti efektif dalam berbagai tugas klasifikasi teks [7][8]. SVM dipilih dalam penelitian ini karena mampu menangani data berdimensi tinggi, cocok untuk pola komentar suporter yang singkat namun sarat fitur emosional, dan memiliki performa yang baik pada tugas klasifikasi teks [6][8]. Untuk merepresentasikan data teks ke dalam bentuk numerik, metode Term Frequency-Inverse Document Frequency (TF-IDF) digunakan karena mampu mengukur tingkat kepentingan suatu kata dalam dokumen secara relatif terhadap keseluruhan korpus [9][10]. Kombinasi TF-IDF dan SVM telah menunjukkan hasil yang baik dalam sejumlah penelitian analisis sentimen di berbagai domain [6][7].

Penelitian terdahulu oleh Jon [2] telah melakukan analisis sentimen pada Instagram Persija Jakarta menggunakan metode Naive Bayes, namun belum mengeksplorasi penggunaan SVM pada konteks yang sama. Irwanda dkk. [3] memang telah menggunakan SVM dalam analisis sentimen, namun penelitian tersebut berfokus pada sentimen masyarakat terhadap Liga Indonesia secara umum, bukan pada klub sepak bola tertentu. Novelty penelitian ini mencakup tiga aspek: (1) penerapan SVM pertama pada konteks Instagram Persija Jakarta, (2) penggunaan data dari dua kondisi pertandingan berbeda untuk menangkap dinamika sentimen yang lebih representatif, dan (3) penyediaan baseline performa SVM pada domain komunitas suporter sepak bola Indonesia.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk mengukur tingkat akurasi SVM dengan kernel linear dan TF-IDF dalam mengklasifikasikan sentimen komentar Instagram Persija Jakarta ke dalam kategori positif, negatif, dan netral, serta memberikan kontribusi pada pengembangan analisis sentimen di domain olahraga.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan Cross Industry Standard Process for Data Mining (CRISP-DM) sebagai kerangka kerja dalam proses analisis sentimen. CRISP-DM terdiri dari enam tahapan sistematis yang saling berkaitan, yaitu Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment.



Gambar 1. Metode penelitian CRISP-DM
(Sumber : https://medium.com)

A. *Business Understanding*

Tahap ini merupakan tahap awal yang bertujuan untuk memahami tujuan dan kebutuhan penelitian. Permasalahan yang diidentifikasi adalah tingginya volume komentar pada akun Instagram resmi Persija Jakarta yang tidak memungkinkan untuk dianalisis secara manual. Oleh karena itu, dibutuhkan sistem otomatis yang mampu mengklasifikasikan sentimen komentar ke dalam tiga kategori, yaitu positif, negatif, dan netral.

B. *Data Understanding*

Pada tahap ini dilakukan proses pengumpulan dan pemahaman terhadap data yang akan digunakan. Data berupa komentar pengguna dikumpulkan dari akun Instagram resmi Persija Jakarta menggunakan layanan Export Comments (https://exportcomments.com/). Data dikumpulkan dari 60 postingan, terdiri dari 30 postingan saat Persija Jakarta menang dan 30 postingan saat kalah, dalam rentang waktu 1 Juli 2023 hingga 23 Mei 2026, sehingga diperoleh total 6.000 komentar.

C. *Data Preparation*

Tahap persiapan data merupakan tahap terpanjang dalam proses CRISP-DM. Pada tahap ini dilakukan serangkaian proses preprocessing untuk mengubah data teks mentah menjadi data yang siap digunakan dalam pemodelan. Tahapan preprocessing yang diterapkan adalah sebagai berikut: (1) *Case Folding*: mengubah seluruh teks komentar menjadi huruf kecil; (2) *Cleaning*: menghapus karakter yang tidak relevan; (3) Normalisasi: mengubah kata tidak baku menjadi bentuk baku; (4) Tokenisasi: memecah teks komentar menjadi satuan kata; (5) *Stopword Removal*: menghapus kata-kata yang tidak memiliki makna signifikan; (6) *Stemming*: mengubah kata berimbuhan menjadi kata dasar; (7) Labeling: pemberian label sentimen secara otomatis menggunakan kamus *INSET Lexicon.*

Setelah proses preprocessing selesai, diperoleh 4.477 komentar yang valid untuk digunakan dalam penelitian. Selanjutnya dilakukan ekstraksi fitur menggunakan metode TF-IDF dengan library Scikit-learn, menghasilkan 4.440 fitur yang merepresentasikan seluruh kosakata dalam dataset. Ekstraksi fitur dilakukan menggunakan metode TF-IDF dengan bantuan library Scikit-learn (`TfidfVectorizer`) menggunakan parameter default. Proses perhitungan TF-IDF terdiri dari tiga tahap, yaitu perhitungan *Term Frequency* (TF), *Inverse Document Frequency* (IDF), dan normalisasi vektor. *Term Frequency* dihitung berdasarkan jumlah kemunculan mentah (*raw count*) suatu term pada dokumen, sebagaimana persamaan (1):

$$TF(t,d) = ft,d \quad (1)$$

dengan $f_{t,d}$ adalah jumlah kemunculan term *t* pada dokumen *d*.

Selanjutnya, Inverse Document Frequency dihitung menggunakan persamaan (2):

$$IDF(t) = \log\left(\frac{1+N}{1+df(t)}\right) + 1 \quad (2)$$

Dengan N adalah jumlah total dokumen dalam korpus, dan *df* (t) adalah jumlah dokumen yang mengandung term t. Penambahan konstanta 1 pada pembilang, penyebut, dan hasil akhir (*smoothing*) bertujuan untuk mencegah nilai IDF menjadi nol maupun pembagian dengan nol, sebagaimana implementasi default pada *Scikit-learn*.

Nilai bobot TF-IDF diperoleh dari perkalian antara TF dan IDF, sebagaimana persamaan (3):

$$TF - IDF(t,d) = TF(t,d) \times IDF(t) \quad (3)$$

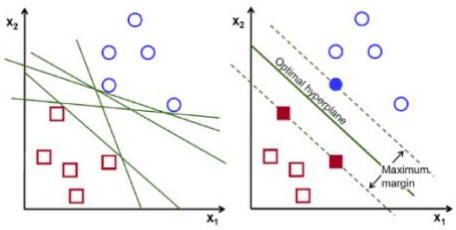
Vektor bobot TF-IDF setiap dokumen kemudian dinormalisasi menggunakan normalisasi L2 agar seluruh vektor memiliki panjang (norma) yang seragam, sebagaimana persamaan (4):

$$V_{norm} = \frac{V}{\|V\|_2} \quad (4)$$

Dengan V adalah vektor bobot TF-IDF suatu dokumen sebelum dinormalisasi, dan $\|V\|_2$ adalah norma Euclidean (L2 norm) dari vektor tersebut. Melalui proses ini, diperoleh 4.440 fitur yang merepresentasikan seluruh kosakata dalam dataset.

D. *Modeling*

Pada tahap ini dilakukan pembangunan model klasifikasi menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel linear SVM bekerja dengan mencari *hyperplane* optimal yang memisahkan antar kelas data dengan margin terbesar, seperti diilustrasikan pada Gambar 2.



Gambar 2. Ilustrasi *Hyperplane* Optimal SVM dengan Maximum Margin **(sumber : https://codingstudio.id/blog/support-vector-machine/)**

Secara matematis, SVM *linear* bertujuan mencari *hyperplane* pemisah dengan persamaan (5):

$$w \cdot x + b = 0 \quad (5)$$

dengan w adalah vektor bobot (normal *hyperplane*), x adalah vektor fitur input, dan b adalah bias.

Klasifikasi data baru dilakukan menggunakan fungsi keputusan pada persamaan (6):

$$f(x) = sign(w \cdot x + b) \quad (6)$$

dengan fungsi sign(.) menghasilkan nilai +1 jika hasil perhitungan w·x+b bernilai positif, dan -1 jika bernilai negatif, yang masing-masing merepresentasikan kelas hasil klasifikasi. Pada penelitian ini, model dilatih menggunakan library *Scikit-learn* (SVC) dengan *kernel linear* untuk menentukan nilai w dan b yang optimal berdasarkan data latih. Dataset dibagi menjadi data latih dan data uji menggunakan metode *train-test split* dengan rasio 80:20, sehingga diperoleh 3.581 data latih dan 896 data uji. Model SVM dilatih menggunakan data latih dan selanjutnya digunakan untuk memprediksi sentimen pada data uji menggunakan *library Scikit-learn*.

E. *Evaluation*

Tahap evaluasi dilakukan untuk mengukur performa model SVM yang telah dibangun menggunakan metrik *accuracy, precision, recall*, dan *F1-score*, serta *confusion matrix*. Hasil evaluasi menunjukkan bahwa model SVM dengan kernel linear mencapai akurasi sebesar 93,64%. Visualisasi hasil evaluasi ditampilkan menggunakan *library Matplotlib* dan *Seaborn*, serta *WordCloud* untuk menampilkan kata dominan pada masing-masing kategori sentimen.

F. *Deployment*

Pada tahap ini, hasil analisis sentimen divisualisasikan menggunakan *framework Streamlit* yang dijalankan secara lokal. Visualisasi yang ditampilkan meliputi distribusi kelas sentimen, *confusion matrix*, serta *WordCloud* dari kata-kata yang paling dominan muncul pada masing-masing kategori sentimen.

III. HASIL DAN PEMBAHASAN

A. *Hasil Pengumpulan Data*

Proses pengumpulan data dilakukan menggunakan layanan *Export Comments* pada akun Instagram resmi Persija Jakarta. Data dikumpulkan dari 60 postingan dalam rentang waktu 1 Juli 2023 hingga 23 Mei 2026, terdiri dari 30 postingan saat Persija menang dan 30 postingan saat kalah, sehingga diperoleh total 6.000 komentar. Data kemudian disimpan dalam format *Excel* (.xlsx) untuk diproses pada tahap selanjutnya.

B. *Hasil Preprocessing*

Data mentah sebanyak 6.000 komentar diproses melalui serangkaian tahapan *preprocessing*. Setelah *preprocessing* selesai, diperoleh 4.477 komentar valid yang siap digunakan dalam penelitian. Sebanyak 1.523 komentar dieliminasi karena merupakan duplikasi, komentar kosong, atau komentar yang tidak mengandung kata bermakna setelah proses *cleaning* dan *stopword removal*. Tabel

menunjukkan contoh hasil preprocessing pada beberapa komentar.

Tabel 1. Contoh Hasil Preprocessing

No	Komentar Asli	Hasil Preprocessing
1	Semangat Persija!!! 🔥🔥 Kamu pasti bisa!!	semangat persija
2	@user123 wkwkwk kalah lagi mah biasa 😂😂	wkwk kalah biasa
3	Wasit nya gak adil banget sih, kecewa banget	wasit adil kecewa
4	Mantap banget permainannya, dukung terus Persija!	mantap main dukung persija
5	udh bayar mahal2 masa kalah mulu sih persija 😡	bayar mahal kalah persija

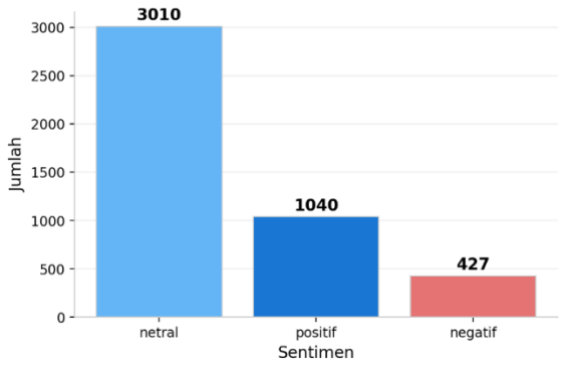
C. *Hasil Pelabelan Data*

Pelabelan sentimen dilakukan secara otomatis menggunakan kamus sentimen Bahasa Indonesia *INSET Lexicon*, kemudian divalidasi secara manual. Hasil pelabelan menunjukkan distribusi kelas yang tidak seimbang (imbalanced), dengan kelas netral mendominasi dataset. Distribusi label sentimen ditunjukkan pada Tabel 2

Tabel 2. Distribusi Label Sentimen

Sentimen	Jumlah	Presentase
Netral	3.010	67,23%
Positif	1.040	23,23%
Negatif	427	9,54%
Total	4.477	100%

Berdasarkan Tabel 2, kelas netral mendominasi dengan 3.010 komentar (67,23%), diikuti kelas positif sebanyak 1.040 komentar (23,23%), dan kelas negatif sebanyak 427 komentar (9,54%). Dominasi komentar netral menunjukkan bahwa sebagian besar pendukung Persija Jakarta cenderung memberikan komentar yang bersifat informatif atau tidak mengandung muatan emosional yang kuat.



Gambar 3. Grafik Batang Distribusi Sentimen

D. *Hasil Ekstraksi Fitur TF-IDF*

Proses ekstraksi fitur dilakukan menggunakan metode TF-IDF terhadap 4.477 komentar yang telah diproses. Hasil ekstraksi menghasilkan 4.440 fitur yang merepresentasikan seluruh kosakata unik dalam dataset. Dataset kemudian dibagi menggunakan metode *train-test split* dengan rasio 80:20, menghasilkan 3.581 data latih dan 896 data uji.

E. *Hasil Pelatihan dan Evaluasi Model SVM*

Model SVM dengan kernel linear dilatih menggunakan 3.581 data latih dan diuji pada 896 data uji. Hasil evaluasi model ditunjukkan pada Gambar 4.

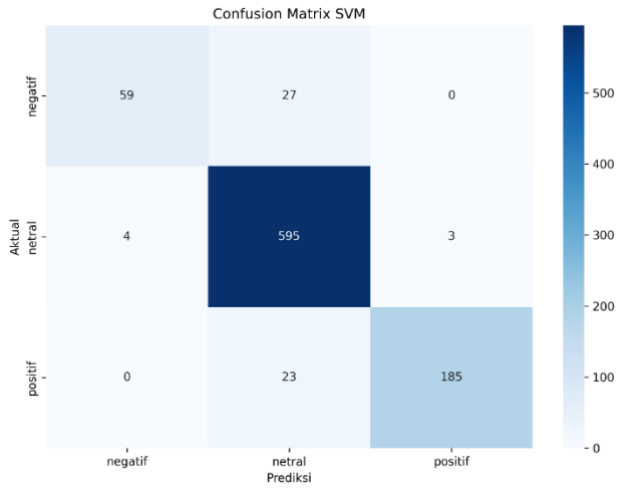
```
==================================
HASIL EVALUASI
==================================
Accuracy : 93.64%

Classification Report:

              precision    recall  f1-score   support

     negatif       0.94      0.69      0.79        86
      netral       0.92      0.99      0.95       602
     positif       0.98      0.89      0.93       208

    accuracy                           0.94       896
   macro avg       0.95      0.85      0.89       896
weighted avg       0.94      0.94      0.93       896
```

Gambar 4. Clasification Report Model SVM

Berdasarkan Gambar 4, model SVM dengan kernel linear mencapai akurasi sebesar 93,64%. Nilai precision tertinggi diperoleh pada kelas positif sebesar 98%, yang berarti dari seluruh komentar yang diprediksi positif, 98% di antaranya benar-benar positif. Nilai recall tertinggi diperoleh pada kelas netral sebesar 99%, yang berarti model mampu

mengenali hampir seluruh komentar netral dengan sangat baik. Nilai F1-score tertinggi juga diperoleh pada kelas netral sebesar 0,95.

Kelas negatif memperoleh recall terendah sebesar 0,69, yang berarti terdapat 31% komentar negatif yang tidak berhasil dikenali oleh model. Hal ini dapat disebabkan oleh jumlah data kelas negatif yang relatif lebih sedikit dibandingkan kelas lainnya (427 data), sehingga model kurang optimal dalam mempelajari pola komentar negatif. Berdasarkan confusion matrix, dari 86 data uji kelas negatif, sebanyak 59 komentar berhasil diklasifikasikan dengan benar, 27 komentar diklasifikasikan sebagai netral, dan tidak ada komentar negatif yang salah diklasifikasikan sebagai positif.



Gambar 5. Confusion Matrix

F. Hasil Visualisasi WordCloud

Visualisasi *WordCloud* digunakan untuk menampilkan kata-kata yang paling dominan muncul pada masing-masing kategori sentimen. Berdasarkan WordCloud pada Gambar 6, kata-kata yang dominan pada kelas positif antara lain "menang", "semangat", "juara", "mantap", "bagus", dan "dukung", yang mencerminkan ekspresi dukungan dan kebanggaan suporter terhadap tim. Pada kelas netral (Gambar 7), kata dominan meliputi "persija", "main", "lawan", dan "liga", yang bersifat informatif. Sementara pada kelas negatif (Gambar 8), kata-kata seperti "kalah", "kecewa", "wasit", "payah", dan "malu" mendominasi, mencerminkan ekspresi kekecewaan dan kritik suporter.

Gambar 4: WordCloud Sentimen Positif

Gambar 5: WordCloud Sentimen Netral

Gambar 6: WordCloud Sentimen Negatif

IV. KESIMPULAN

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa penelitian ini sukses mengklasifikasikan sentimen komentar pada akun Instagram resmi Persija Jakarta ke dalam kategori positif, negatif, dan netral. Dengan menerapkan metode *Support Vector Machine* (SVM) kernel linear yang dikombinasikan dengan ekstraksi fitur *Term Frequency-Inverse Document Frequency* (TF-IDF), model yang dibangun terbukti sangat efektif. Hal ini ditunjukkan oleh capaian akurasi yang tinggi, yaitu sebesar 93,64%, dengan tingkat *precision* tertinggi berada pada kelas positif sebesar 98%, serta *recall* tertinggi pada kelas netral yang mencapai 99%.

Keberhasilan ini tidak lepas dari tahapan preprocessing teks dan proses pelabelan

menggunakan *INSET Lexicon* yang berhasil menyaring 6.000 komentar mentah hingga menghasilkan 4.477 data yang valid. Dari dataset valid tersebut, proses ekstraksi fitur TF-IDF mampu menghasilkan 4.440 representasi kosakata unik. Seluruh data ini kemudian dibagi menggunakan rasio 80:20, yang menghasilkan 3.581 data latih dan 896 data uji untuk proses pemodelan.

Dari hasil analisis data tersebut, ditemukan bahwa distribusi sentimen suporter didominasi oleh komentar netral sebesar 67,23%, disusul oleh komentar positif sebesar 23,23%, dan komentar negatif sebesar 9,54%. Dominasi angka ini mengindikasikan bahwa sebagian besar interaksi suporter Persija Jakarta di platform Instagram cenderung bersifat informatif dan objektif, alih-alih berupa ekspresi emosional yang meluap-luap.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan dan kontribusi dalam penyelesaian penelitian ini. Ucapan terima kasih yang sebesar-besarnya disampaikan kepada dosen pembimbing yang telah memberikan bimbingan, arahan, dan masukan yang sangat berarti selama proses penelitian berlangsung. Penulis juga mengucapkan terima kasih kepada program studi dan institusi yang telah menyediakan fasilitas dan dukungan akademik sehingga penelitian ini dapat terlaksana dengan baik. Tidak lupa, penulis menyampaikan apresiasi kepada seluruh pihak yang telah membantu dalam proses pengumpulan data maupun pengujian sistem. Semoga penelitian ini dapat memberikan manfaat dan kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan pembelajaran mesin berbahasa Indonesia.

REFERENSI

GoodStats. (2025). Simak tren pengguna Instagram di Indonesia 2018-2025. https://data.goodstats.id/chart/pengguna-instagram-di-indonesia

Jon, A. M. (2023). Analisis Sentimen pada Media Sosial Instagram Klub Persija Jakarta menggunakan Metode Naive Bayes. Universitas Islam Indonesia.

Irwanda, M., Afdal, M., Novita, R., & Zarnelly, Z. (2025). Analisis sentimen masyarakat terhadap Liga Indonesia menggunakan Naive Bayes Classifier dan SVM. BITS, 7(1).

Tambunan, H. B., & Hapsari, T. W. D. (2022). Analisis opini pengguna aplikasi New PLN Mobile menggunakan text mining. PETIR, 15(1), 121-130.

Sirojj, D. E. H. (2023). Analisis Sentimen terhadap Qatar sebagai Tuan Rumah Piala Dunia FIFA 2022 pada Media Sosial Twitter menggunakan Algoritma C4.5. UIN Maulana Malik Ibrahim Malang.

Purwanti, Z., & Sugiyono. (2024). Pemodelan text mining untuk analisis sentimen menggunakan SVM. Jurnal Indonesia: Manajemen Informatika dan Komunikasi, 5(3), 3065-3079.

Gifari, O. I., Adha, M., Hendrawan, I. R., & Durrand, F. F. S. (2022). Analisis sentimen review film menggunakan TF-IDF dan Support Vector Machine. JIFOTECH, 2(1), 36-45.

Aisah, I. S., Irawan, B., & Suprapti, T. (2023). Algoritma Support Vector Machine (SVM) untuk analisis sentimen ulasan aplikasi Al Qur'an digital. JATI: Jurnal Mahasiswa Teknik Informatika, 7(6), 3759-3768.

Pratama, R. A., et al. (2024). Analisis Sentimen Ulasan Aplikasi menggunakan KNN dengan TF-IDF.

Huda, A. A., Fajarudin, R., & Hadinegoro, A. (2022). Sistem rekomendasi content-based filtering menggunakan TF-IDF. BITS, 4(3), 1679-1686.

Sutaryani, A., Sunarno, S., & Djuniadi, D. (2023). Perbandingan performa model machine learning dalam prediksi suhu di Semarang. JITET, 12(3).

Fahmi, M. N. (2023). Implementasi machine learning menggunakan Python library: Scikit-learn. Sains Data, 1(2), 87-96.

Ardiansyah, Widagdo, A. S., Qodri, K. N., Saputro, F. E. N., & Rizky, N. A. (2025). Penerapan Lexicon Based untuk Analisis Sentimen Masyarakat Indonesia terhadap Danantara. JATI: Jurnal Mahasiswa Teknik Informatika, 9(3), 4949-4957.

Mudya, A., et al. (2024). Analisis Sentimen Menggunakan Kamus Lexicon. VARIANSI: Journal of Statistics and Its Application on Teaching and Research, 6(2), 76-83.

Huda, A. A., Fajarudin, R., & Hadinegoro, A. (2022). Sistem rekomendasi content-based filtering menggunakan TF-IDF. BITS, 4(3), 1679-1686.