Evaluasi Kepuasan Penggemar Sepak Bola Terhadap Pemilihan Pelatih Timnas Indonesia Di Media Sosial X Dengan Metode K-Means Clustering

Evaluation of Football Fans' Satisfaction with the Selection of Indonesia's National Team Coach on Social Media X Using the K-Means Clustering Method

Nasywa Al Afif Harahap[1], **Abdul Halim Hasugian**[2]
Fakultas Sains dan Teknologi, Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
E-mail : afifhrp21@gmail.com[1]*, abdulhalimhasugian@uinsu.ac.id*[2]
Received 5 July 2025; Revised 29 July 2025; Accepted 8 August 2025

Abstrak - Tingginya antusiasme publik terhadap pemilihan pelatih timnas Indonesia seringkali memunculkan beragam opini di media sosial, khususnya platform X. Opini tersebut tersebar dalam bentuk komentar yang tidak terstruktur, sehingga menyulitkan evaluasi kepuasan publik secara objektif. Penelitian ini merumuskan permasalahan: bagaimana mengelompokkan opini publik terhadap pemilihan pelatih timnas Indonesia secara sistematis untuk mengevaluasi tingkat kepuasan penggemar. Tujuan penelitian ini adalah menerapkan algoritma K-Means Clustering dalam proses analisis sentimen berbasis teks untuk mengetahui persepsi publik secara terukur. Penelitian ini menggunakan pendekatan kuantitatif dengan tahapan utama berupa crawling data tweet, text preprocessing, pembobotan TF-IDF, serta klasterisasi menggunakan metode K-Means. Penentuan jumlah klaster optimal dilakukan dengan Elbow Method dan validasi menggunakan Silhouette Score. Hasil analisis terhadap 947 data menunjukkan distribusi sentimen positif sebanyak 649 tweet (68,46%), netral 185 tweet (19,51%), dan negatif 114 tweet (12,03%). Evaluasi performa menghasilkan akurasi model sebesar 53,59%, dengan performa terbaik pada klaster sentimen positif. Penelitian menyimpulkan bahwa metode K-Means Clustering dapat menjadi pendekatan awal dalam menganalisis opini publik di media sosial, meskipun akurasinya masih terbatas untuk data dengan distribusi tidak seimbang. Penelitian ini bermanfaat dalam memberikan rekomendasi berbasis data bagi federasi sepak bola Indonesia untuk memahami suara publik sebagai bahan evaluasi dalam pengambilan keputusan strategis.
Kata kunci - Analisis Sentimen, K-Means Clustering, Machine Learning, TF-IDF, Confusion Matrix

Abstract *- The high public enthusiasm for the selection of the Indonesian national football team coach often sparks a wide range of opinions on social media, particularly on the X platform. These opinions are scattered in the form of unstructured comments, making it difficult to objectively assess public satisfaction. This study formulates the problem: how to systematically categorize public opinion regarding the selection of the national team coach to evaluate fan satisfaction levels. The aim of this research is to apply the K-Means Clustering algorithm in a text-based sentiment analysis process to quantitatively measure public perception. A quantitative approach is used, with key stages including tweet data crawling, text preprocessing, TF-IDF weighting, and clustering using the K-Means method. The optimal number of clusters is determined using the Elbow Method and validated with the Silhouette Score. Analysis of 947 data points revealed sentiment distribution as follows: 649 positive tweets (68.46%), 185 neutral tweets (19.51%), and 114 negative tweets (12.03%). Model performance evaluation yielded an accuracy of 53.59%, with the best performance observed in the positive sentiment cluster. The study concludes that the K-Means Clustering method can serve as an initial approach to analyzing public opinion on social media, although its accuracy remains limited for data with imbalanced distribution. This research is valuable in providing data-driven recommendations for*

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

the Indonesian football federation to better understand public sentiment as input for strategic decision-making.
Keywords *— Sentiment Analysis, K-Means Clustering, Machine Learning, TF-IDF, Confusion Matrix*

1. PENDAHULUAN

Perkembangan teknologi informasi yang semakin pesat telah membawa perubahan signifikan di berbagai aspek kehidupan, termasuk dalam dunia olahraga. Teknologi informasi tidak hanya digunakan untuk mengolah dan menyimpan data, tetapi juga menjadi sarana komunikasi dan interaksi sosial yang sangat penting dalam kehidupan modern [1]. Salah satu cabang olahraga yang paling populer di Indonesia, yaitu sepak bola, juga turut merasakan dampak dari kemajuan teknologi ini. Sepak bola tidak hanya menjadi tontonan hiburan, tetapi telah berkembang menjadi budaya yang memiliki basis penggemar yang sangat besar dan loyal. Setiap keputusan penting yang diambil dalam dunia sepak bola, seperti pemilihan pelatih tim nasional, selalu menjadi perbincangan hangat di tengah masyarakat.

Media sosial, khususnya platform X, telah menjadi wadah utama bagi para penggemar sepak bola untuk menyampaikan opini, kritik, dan dukungan mereka terhadap kebijakan yang dikeluarkan oleh federasi sepak bola Indonesia. Dinamika opini yang muncul di platform ini mencerminkan bagaimana persepsi publik terhadap keputusan yang diambil oleh otoritas sepak bola, termasuk dalam hal pemilihan pelatih tim nasional. Namun, opini-opini tersebut seringkali sulit diukur secara kuantitatif dan tidak dapat langsung disimpulkan maknanya tanpa proses analisis yang sistematis.

X sebagai salah satu media sosial berbasis teks memiliki karakteristik unik, di mana penggunanya dapat secara langsung menyuarakan pendapat dan pandangan mereka dalam bentuk komentar, balasan, atau unggahan yang sering kali dipicu oleh topik-topik trending. Keberagaman opini yang terkandung dalam komentar-komentar ini berpotensi besar untuk dianalisis lebih lanjut dalam rangka memahami kepuasan penggemar terhadap kebijakan federasi sepak bola Indonesia. Dalam konteks inilah, metode analisis data seperti clustering dapat memainkan peran penting untuk mengelompokkan opini berdasarkan tema dan sentimen yang terkandung di dalamnya [2].

Penelitian ini menggunakan metode K-Means Clustering sebagai pendekatan utama dalam proses analisis [3]. Metode ini dipilih karena memiliki beberapa keunggulan yang relevan dengan karakteristik data opini publik di media sosial. K-Means merupakan metode unsupervised learning yang tidak membutuhkan label sentimen secara eksplisit, sehingga cocok untuk data mentah dari media sosial [4]. Selain itu, metode ini mampu mengelompokkan data berdasarkan kemiripan konten menggunakan perhitungan jarak antar data, serta efisien untuk data berukuran besar seperti kumpulan tweet. Dengan penerapan K-Means Clustering, diharapkan opini-opini publik dapat dikelompokkan secara sistematis ke dalam beberapa klaster yang mencerminkan tingkat kepuasan atau ketidakpuasan terhadap pemilihan pelatih baru tim nasional Indonesia [5].

Permasalahan yang diangkat dalam penelitian ini berangkat dari tingginya atensi publik terhadap keputusan strategis yang diambil oleh federasi sepak bola Indonesia, khususnya terkait pemilihan pelatih tim nasional. Antusiasme dan ekspektasi penggemar terhadap performa timnas kerap kali diwujudkan dalam bentuk opini yang disampaikan secara terbuka di media sosial. Namun, opini-opini tersebut umumnya bersifat tidak terstruktur dan tersebar, sehingga menyulitkan proses analisis secara sistematis. Oleh karena itu, penelitian ini difokuskan untuk memahami sejauh mana kepuasan penggemar sepak bola Indonesia terhadap keputusan pemilihan pelatih baru yang tercermin melalui opini mereka di platform media sosial X. Berbagai studi terdahulu telah banyak membahas penerapan algoritma K-Means dalam berbagai domain seperti pendidikan, layanan kesehatan, dan pariwisata. Namun, penelitian terkait opini penggemar sepak bola terhadap pemilihan pelatih tim nasional Indonesia di platform X masih sangat terbatas. Hal ini menunjukkan adanya celah penelitian yang ingin dijembatani oleh studi ini.

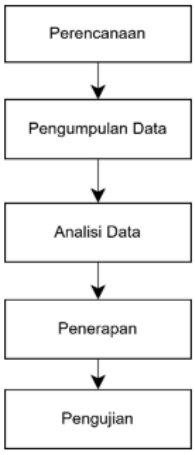
e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

Penelitian ini didukung oleh berbagai studi terdahulu seperti yang dilakukan oleh Paembonan dan Abduh (2021) dimana menerapkan K-Means dalam pengelompokan obat, memperoleh nilai silhouette coefficient 0,4854 yang menunjukkan Metode ini membantu dalam mengelompokkan obat dengan karakteristik serupa, sehingga memudahkan proses rekomendasi obat pengganti bagi pasien [6]. Penelitian yang dilakukan oleh Sholeh dan Aeni (2023) mengkaji pengelompokan data ulasan destinasi wisata menggunakan algoritma K-Means dengan pendekatan Knowledge Discovery in Database (KDD). Proses penelitian mencakup tahapan seleksi data, pembersihan, transformasi, hingga evaluasi model dengan metode Davies Bouldin, Elbow, dan Silhouette. Ketiga metode tersebut merekomendasikan jumlah klaster optimal sebanyak dua, yang mewakili kategori puas dan tidak puas. Hal ini menunjukkan bahwa kombinasi algoritma K-Means dan evaluasi berbasis metrik klasterisasi dapat memberikan hasil yang akurat dalam pengelompokan opini publik [7].

Sementara itu, penelitian oleh Rheza dan Metandi (2020) menunjukkan bahwa algoritma K-Means juga dapat dimanfaatkan untuk mengelompokkan komentar berdasarkan gaya bahasa. Dalam studi tersebut, 50 komentar X tentang PSSI dikelompokkan menjadi tiga jenis, yaitu ironi, sinisme, dan sarkasme. Hasilnya menunjukkan bahwa komentar ironi mendominasi dengan proporsi 54%, diikuti sinisme 40% dan sarkasme 6%. Penelitian ini membuktikan bahwa metode K-Means efektif dalam mengidentifikasi pola komunikasi yang khas di media sosial, khususnya dalam konteks isu-isu sepak bola nasional [8]. Adapun penelitian oleh Agustian et al. (2024) memadukan algoritma K-Means dengan Naive Bayes untuk memetakan opini publik terhadap kebijakan naturalisasi pemain sepak bola tim nasional Indonesia. Dengan menggunakan 3.584 komentar dari platform YouTube, hasil klasifikasi menunjukkan akurasi sebesar 93,17%, menandakan bahwa kombinasi kedua metode tersebut cukup efektif dalam mengelompokkan opini publik [9]. Melalui penelitian ini, diharapkan diperoleh wawasan yang lebih mendalam mengenai persepsi penggemar terhadap pemilihan pelatih tim nasional Indonesia. Temuan ini dapat menjadi bahan evaluasi dan pertimbangan bagi pemangku kebijakan di bidang sepak bola nasional, agar setiap keputusan yang diambil dapat lebih selaras dengan aspirasi publik.

2. METODE PENELITIAN

Penelitian ini menggunakan metodologi kuantitatif yang menguji teori dengan melihat hubungan antar variabel. Tujuan dari penelitian ini adalah untuk mengembangkan evaluasi kepuasan penggemar sepak bola terhadap pemilihan pelatih baru timnas indonesia di media sosial X dengan metode k means clustering. Terdapat beberapa tahapan yang harus dilalui untuk melakukan penelitian ini, diantaranya sebagai berikut.



Gambar 1. Kerangka Penelitian

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

2.1. Perencanaan Penelitian

Proses penelitian ini diawali dengan perencanaan yaitu menentukan topik yang akan akan dibahas. Topik penelitian ini yaitu implementasi text mining dalam melihat kepuasan penggemar sepak bola terhadap pemilihan pelatih baru timnas indonesia yang sumber datanya berasal dari aplikasi sosial media X menggunakan algoritma K Means Clustering.

2.2. Pengumpulan Data

Pada tahap ini, data berhasil dikumpulkan dengan cara melakukan Crawling Data pada aplikasi X tentang kepuasan para penggemar sepak bola terkait pelatih baru timnas Indonesia dengan keyword #pelatih, #patrickkluivert, #timnas, #kualifikasi, #pemilihan, #timnas senior. Dengan dilakukannya crawling tersebut, berhasil didapatkan sebanyak lebih dari 1000 data dari masing-masing komentar para penggemar, yang akan digunakan dalam proses klasifikasi kepuasan menggunakan metode K Means Clustering. Berikut beberapa sampel data yang berhasil didapatkan, diantaranya :

Tabel 1. Sampel Data

No	Username	Tweet
1	@BillyAditia_	@TimnasIndonesia@PSSI@PatrickKluivert Buset terakhir jadi pelatih kepala 17 tahun lalu.
2	@leotbnid	@TimnasIndonesia @PSSI @PatrickKluivert Mantap ditunggu sbg pelatih apakah mampu membawa Indonesia lebih tinggi.
3	@arif_prnm	@TimnasIndonesia @PSSI @PatrickKluivert Setelah Pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih2 belanda yg di akademi mantap.

2.3. Analisis Data

Tahap analisis data setelah dikumpulkan dari Media Sosial X menggunakan teknik Crawling Data untuk memperoleh komentar, tweet, atau unggahan yang membahas pemilihan pelatih Timnas Indonesia. selanjutnya dimulai dengan preprocessing untuk membersihkan data dari unsur yang tidak perlu, seperti tanda baca, angka, serta kata-kata umum (stopwords). Kemudian, dilakukan proses tokenization, dan stemming agar data lebih terstruktur. Kemudian dilakukan pembobotan dengan TF-IDF sebelum data diklasifikasikan menggunakan algoritma K-means clustering. Setelah itu, menentukan jumlah klaster optimal (K) menggunalan metode Elbow Method. Disini peneliti menentukan 3 klaster yang dikelompokkan berdasarkan klaster 1 (positif), klaster 2 (negatif) dan klaster 3 (netral).

2.4. Penerapan K-Means Clustering

K-Means Clustering merupakan salah satu algoritma unsupervised learning yang umum digunakan untuk mengelompokkan data ke dalam sejumlah kelompok atau klaster berdasarkan kemiripan karakteristik. Metode ini bekerja dengan mempartisi data ke dalam k klaster, di mana masing-masing klaster memiliki pusat (centroid) yang menjadi acuan pengelompokan [10]. Tujuan dari algoritma ini adalah untuk meminimalkan jarak antar data dalam satu klaster dan memaksimalkan jarak antar klaster [11]. Proses K-Means dimulai dengan menentukan jumlah klaster (k) dan memilih pusat klaster secara acak. Selanjutnya, jarak antara setiap data dan centroid dihitung menggunakan rumus Euclidean Distance [12]:

$$d(x_{i,}\mu_j) = \sqrt{\sum(xi - \mu j)^2} \qquad (1)$$

Rumus tersebut menghitung jarak antara data ke-i (xi) dan centroid klaster ke-j (μj) berdasarkan nilai atribut yang dimiliki masing-masing data. Setelah jarak dihitung, setiap data

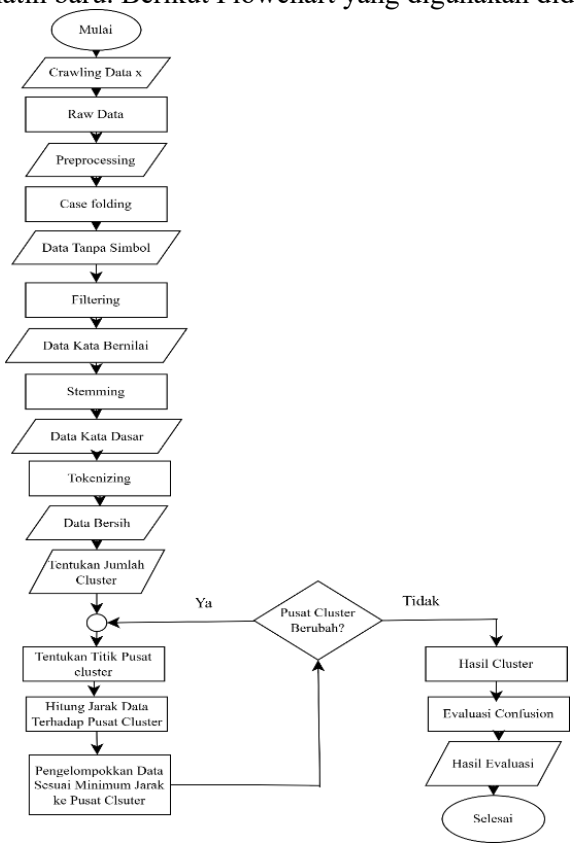
e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

akan dikelompokkan ke klaster dengan centroid terdekat. Kemudian, nilai centroid diperbarui dengan menghitung rata-rata dari seluruh data dalam klaster tersebut menggunakan rumus:

$$C_j = \frac{1}{N_k} \sum_{k=1}^{N_k} x_i \qquad (2)$$

Pada rumus di atas, Cj merupakan centroid baru untuk klaster ke-j, Nk adalah jumlah data dalam klaster tersebut, dan xi adalah data individual di dalamnya. Proses ini dilakukan secara iteratif, yaitu penghitungan jarak, pengelompokan ulang, dan pembaruan centroid, hingga tidak terjadi lagi perubahan signifikan pada posisi centroid. Ketika kondisi tersebut tercapai, maka proses klasterisasi dianggap selesai.

Dalam penelitian ini, algoritma K-Means digunakan setelah dilakukan pembobotan kata menggunakan metode TF-IDF terhadap data opini penggemar sepak bola di media sosial X. Tujuannya adalah untuk mengelompokkan opini publik ke dalam klaster tertentu berdasarkan kemiripan konten, sehingga dapat diidentifikasi pola sentimen dan tingkat kepuasan terhadap pelatih baru tim nasional Indonesia. Setelah dilakukannya pembobotan kata dengan TF-IDF, selanjutnya data diklasifikasikan menggunakan metode K Means dalam evaluasi kepuasan para penggemar terhadap pelatih baru. Berikut Flowchart yang digunakan didalam penelitian ini:



Gambar 2. Flowchart K-Means Clustering

2.5. Pengujian

Tahap pengujian pada penelitian evaluasi kepuasan ini menggunakan metode Elbow Method dan Silhouette Score [13] yang nantinya untuk mendapatkan klaster yang optimal dan dilakukan validasi menggunakan split data (training & testing set) atau cross-validation untuk melihat apakah pola klaster tetap konsisten saat diterapkan pada dataset baru.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

3. HASIL DAN PEMBAHASAN

Data dalam penelitian ini diperoleh dari media sosial X menggunakan teknik web scraping melalui Google Colaboratory dengan bahasa pemrograman Python. Proses pengumpulan dilakukan dengan kata kunci seperti #pelatih, #patrickkluivert, #timnas, #kualifikasi, #pemilihan, dan #timnassenior, yang berkaitan dengan opini publik terhadap pemilihan pelatih timnas Indonesia. Hasil scraping menghasilkan lebih dari 1.000 komentar penggemar dalam bentuk teks, yang kemudian diproses melalui tahapan praproses dan pembobotan menggunakan TF-IDF [14]. Data ini selanjutnya dianalisis menggunakan metode K-Means Clustering untuk mengelompokkan opini ke dalam klaster positif, negatif, dan netral, sehingga memberikan gambaran tingkat kepuasan publik secara sistematis.

Tabel 2. Dataset

Dataset
@TimnasIndonesia@PSSI@PatrickKluivert Buset terakhir jadi pelatih kepala 17 tahun lalu.
@TimnasIndonesia @PSSI @PatrickKluivert Mantap ditunggu sbg pelatih apakah mampu membawa Indonesia lebih tinggi.
@TimnasIndonesia @PSSI @PatrickKluivert Setelah Pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih2 belanda yg di akademi mantap.

Proses analisis sentimen kepuasan penggemar sepak bola terhadap pemilihan pelatih timnas Indonesia diawali dengan tahap krusial, yaitu preprocessing data. Data mentah hasil scraping dari media sosial X perlu dibersihkan dari elemen-elemen yang tidak relevan agar siap dianalisis secara komputasional. Tahapan preprocessing yang dilakukan meliputi cleaning, case folding, tokenizing, dan stemming. Pada tahap cleaning, dilakukan penghapusan karakter seperti angka, simbol, emoji, tautan, hashtag, serta tanda baca yang tidak berkontribusi terhadap analisis semantik.

Tabel 3. Proses Cleaning

Sentimen Awal	Cleaning
@TimnasIndonesia@PSSI@PatrickKluivert Buset terakhir jadi pelatih kepala 17 tahun lalu.	Buset terakhir jadi pelatih kepala tahun lalu
@TimnasIndonesia @PSSI @PatrickKluivert Mantap ditunggu sbg pelatih apakah mampu membawa Indonesia lebih tinggi.	Mantap ditunggu sbg pelatih apakah mampu membawa Indonesia lebih tinggi
@TimnasIndonesia @PSSI @PatrickKluivert Setelah Pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih2 belanda yg di akademi mantap.	Setelah Pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih belanda yg di akademi mantap

Langkah selanjutnya adalah tahap case folding. Tahapan ini merupakan tahapan proses konversi perubahan huruf kapital menjadi huruf kecil (lowercase) pada semua sentimen yang ada. Berikut ini adalah merupakan hasil case folding yang ditunjukkan pada tabel 4.

Tabel 4. Sampel Proses Case Folding

Cleaning	Case Folding
Buset terakhir jadi pelatih kepala tahun lalu	buset terakhir jadi pelatih kepala tahun lalu
Mantap ditunggu sbg pelatih apakah mampu membawa Indonesia lebih tinggi	mantap ditunggu sbg pelatih apakah mampu membawa indonesia lebih tinggi
Setelah Pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih belanda yg di akademi mantap	setelah pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih belanda yg di akademi mantap

Setelah proses case folding, dilakukan tahap normalisasi untuk mengganti kata-kata slang atau tidak baku, seperti "gue" menjadi "saya" dan "bgt" menjadi "banget", menggunakan kamus

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

slang. Tahap ini penting agar algoritma dapat memahami makna kata dengan lebih akurat dalam analisis sentimen.

Tabel 5. Sampel Proses Normalization

Case Folding	Normalization
buset terakhir jadi pelatih kepala tahun lalu	buset terakhir jadi pelatih kepala tahun lalu
mantap ditunggu sbg pelatih apakah mampu membawa indonesia lebih tinggi	mantap ditunggu sebagai pelatih apakah mampu membawa indonesia lebih tinggi
setelah pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih belanda yg di akademi mantap	mantap ditunggu sebagai pelatih apakah mampu membawa indonesia lebih tinggi

Selanjutnya adalah tokenisasi yang dimana merupakan langkah penting dalam analisis sentimen dan pemrosesan bahasa alami (NLP) secara keseluruhan. Proses ini bagaikan membongkar teks menjadi bagian-bagian kecil, yang disebut "token", agar lebih mudah dipahami dan dianalisis oleh mesin.

Tabel 6. Sampel Proses Tokenizing

Normalization	Tokenizing
buset terakhir jadi pelatih kepala tahun lalu	['buset', 'terakhir', 'jadi', 'pelatih', 'kepala', 'tahun', 'lalu']
mantap ditunggu sebagai pelatih apakah mampu membawa indonesia lebih tinggi	['mantap', 'ditunggu', 'sebagai', 'pelatih', 'apakah', 'mampu', 'membawa', 'indonesia', 'lebih', 'tinggi']
setelah pemain keturunan kita manfaatkan stok pelatih belanda yg melimpah kemudian kita ambil arsitek juga buat bikin stadion dan pelatih belanda yg di akademi mantap	['setelah', 'pemain', 'keturunan', 'kita', 'manfaatkan', 'stok', 'pelatih', 'belanda', 'yang', 'melimpah', 'kemudian', 'kita', 'ambil', 'arsitek', 'juga', 'buat', 'bikin', 'stadion', 'dan', 'pelatih', 'belanda', 'yang', 'di', 'akademi', 'mantap']

Setelah proses tokenizing, dilakukan tahap stopword removal untuk menghapus kata-kata umum seperti "yang", "dan", "di", dan "itu" yang tidak memiliki makna signifikan dalam analisis sentimen. Penghapusan ini membantu meningkatkan efisiensi dan menekankan kata-kata yang lebih relevan.

Tabel 7. Sampel Proses Stopword Removal

Tokenizing	Stepword Removal
['buset', 'terakhir', 'jadi', 'pelatih', 'kepala', 'tahun', 'lalu']	['buset', 'pelatih', 'kepala']
['mantap', 'ditunggu', 'sebagai', 'pelatih', 'apakah', 'mampu', 'membawa', 'indonesia', 'lebih', 'tinggi']	['mantap', 'ditunggu', 'pelatih', 'membawa', 'indonesia']
['setelah', 'pemain', 'keturunan', 'kita', 'manfaatkan', 'stok', 'pelatih', 'belanda', 'yang', 'melimpah', 'kemudian', 'kita', 'ambil', 'arsitek', 'juga', 'buat', 'bikin', 'stadion', 'dan', 'pelatih', 'belanda', 'yang', 'di', 'akademi', 'mantap']	['pemain', 'keturunan', 'manfaatkan', 'stok', 'pelatih', 'belanda', 'melimpah', 'ambil', 'arsitek', 'bikin', 'stadion', 'pelatih', 'belanda', 'akademi', 'mantap']

Setelah proses stopword removal, tahap selanjutnya adalah stemming, yaitu mengubah kata menjadi bentuk dasarnya atau akar kata. Proses ini bertujuan untuk mengurangi variasi kata yang memiliki makna serupa, seperti "berlari", "lari-lari", dan "berlari-lari" menjadi "lari", atau "makanan" dan "dimakan" menjadi "makan". Dengan demikian, stemming membantu mengurangi dimensi data dan meningkatkan akurasi dalam analisis sentimen.

Tabel 8. Sampel Proses Stemming

Stepword Removal	Stemming
['buset', 'pelatih', 'kepala']	buset latih kepala
['mantap', 'ditunggu', 'pelatih', 'membawa', 'indonesia']	mantap tunggu latih bawa Indonesia
['pemain', 'keturunan', 'manfaatkan', 'stok', 'pelatih', 'belanda', 'melimpah', 'ambil', 'arsitek', 'bikin', 'stadion', 'pelatih', 'belanda', 'akademi', 'mantap']	main turun manfaat stok latih belanda limpah ambil arsitek bikin stadion latih belanda akademi mantap
['kluivert', 'prestasi', 'mengajari', 'pelatih', 'lokal', 'kocak']	kluivert prestasi ajar latih lokal kocak

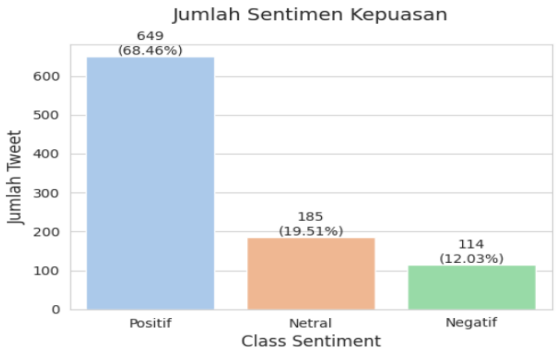
Setelah proses stemming, tahap berikutnya adalah pelabelan sentimen menggunakan VADER Lexicon [15]. VADER digunakan untuk mengidentifikasi polaritas sentimen setiap teks

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

dengan menghasilkan skor dari -1 hingga +1. Dalam penelitian ini, teks dengan skor ≥ 1 dikategorikan sebagai positif, skor ≤ -1 sebagai negatif, dan skor 0 sebagai netral. Pendekatan ini memungkinkan klasifikasi sentimen secara sederhana namun efektif ke dalam tiga kategori utama, sehingga mempermudah proses analisis dan klasterisasi opini pengguna.

Tabel 9. Hasil Pelabelan

Stemming	Sentimen_Score	Label
buset latih kepala	2	Positif
mantap tunggu latih bawa Indonesia	3	Positif
main turun manfaat stok latih belanda limpah ambil arsitek bikin stadion latih belanda akademi mantap	2	Positif



Gambar 3. Pemberian Label Dengan Lexicon Based

Berdasarkan diagram batang di atas yang menggambarkan jumlah sentimen kepuasan, mayoritas tweet yang dianalisis menunjukkan sentimen positif sebanyak 649 tweet (68,46%), yang mencerminkan dominasi persepsi positif dari pengguna terhadap topik yang dibahas. Sentimen netral berada di posisi kedua dengan 185 tweet (19,51%), menandakan sebagian pengguna menyampaikan opini yang tidak condong pada sisi positif maupun negatif. Sementara itu, sentimen negatif berjumlah 114 tweet (12,03%), menunjukkan hanya sebagian kecil pengguna yang mengungkapkan ketidakpuasan.

Dominasi sentimen positif (68,46%) menunjukkan bahwa sebagian besar pengguna X memberikan tanggapan yang optimis terhadap pemilihan pelatih baru. Meskipun demikian, sentimen negatif yang mencapai 12,03% mengindikasikan adanya kelompok yang kurang puas atau skeptis terhadap keputusan tersebut. Sentimen netral yang cukup besar (19,51%) juga menunjukkan bahwa sebagian pengguna memilih bersikap moderat atau menunggu hasil nyata dari pelatih terpilih. Hasil ini mencerminkan kompleksitas opini publik yang perlu diperhatikan oleh pemangku kebijakan.

Setelah pelabelan sentimen, tahap berikutnya adalah pembobotan kata menggunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*). TF-IDF digunakan untuk mengukur seberapa penting sebuah kata dalam suatu dokumen relatif terhadap kumpulan dokumen lainnya. Proses ini diawali dengan menghitung term frequency (TF), yaitu frekuensi kemunculan kata dalam sebuah teks, kemudian dilanjutkan dengan menghitung inverse document frequency (IDF) untuk mengetahui tingkat kekhasan kata dalam seluruh dokumen. Terakhir, bobot kata (weight, W) dihitung dengan mengalikan nilai TF dan IDF. Tahap awal yang dilakukan adalah menghitung nilai TF dari semua kata yang terdapat dalam teks (menghitung seberapa kata muncul). Menggunakan 4 kalimat yaitu:

a. buset latih kepala
b. mantap tunggu latih bawa Indonesia

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

c. main turun manfaat stok latih belanda limpah ambil arsitek bikin stadion latih belanda akademi mantap
d. kluivert prestasi ajar latih lokal kocak

Berikut contoh perhitungan dari proses TF-IDF dari kalimat yang diberikan sebelumnya pada tabel 10.

Tabel 10. Perhitungan nilai TF dan DF

Term	TF				DF
	D1	D2	D3	D4	
Ambil	0	0	1	0	1
Ajar	0	0	0	1	1
Akademi	0	0	1	0	1
Arsitek	0	0	1	0	1
Bawa	0	1	0	0	1
Belanda	0	0	2	0	1
Bikin	0	0	1	0	1
Buset	1	0	0	0	1
Kepala	1	0	0	0	1
Kluivert	0	0	0	1	1
Kocak	0	0	0	1	1
Latih	1	1	2	1	4
Lokal	0	0	0	1	1
Limpah	0	0	1	0	1
Main	0	0	1	0	1
Manfaat	0	0	1	0	1
Mantap	0	1	1	0	2
Prestasi	0	0	0	1	1
Stadion	0	0	1	0	1
Stok	0	0	1	0	1
Turun	0	0	1	0	1
Tunggu	0	1	0	0	1
Indonesia	0	1	0	0	1

Setelah tahap perhitungan TF dan DF dari setiap term selesai dilakukan, selanjutnya tahap perhitungan IDF menggunakan persamaan berikut. Berikut disajikan tabel hasil perhitungan IDF melanjutkan dari tahap perhitungan sebelumnya. Contoh Perhitungan IDF pada term pertama yaitu “ambil” dibawah.

$$Idf_t = \log\frac{4}{1} = 0.602$$

Proses perhitungan nilai IDF tidak dibahas secara kesuluruhan kata untuk hasil nilai IDF disajikan dalam tabel 11.

Tabel 11. Perhitungan Nilai IDF

Term	DF	IDF
Ambil	1	0.602
Ajar	1	0.602
Akademi	1	0.602
…	…	…
Indonesia	1	0.602

Setelah tahap perhitungan nilai IDF dari setiap term selesai dilakukan, selanjutnya tahap perhitungan nilai TF-IDF, berikut disajikan Tabel 12 hasil perhitungan TF-IDF melanjutkan dari tahap perhitungan sebelumnya.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

Tabel 12. Perhitungan Nilai TF IDF

Term	TF-IDF			
	D1	D2	D3	D4
Ambil	0	0	0.602	0
Ajar	0	0	0	0.602
Akademi	0	0	0.602	0
Arsitek	0	0	0.602	0
Bawa	0	0.602	0	0
Belanda	0	0	1.204	0
Bikin	0	0	0.602	0
Buset	0.602	0	0	0
Kepala	0.602	0	0	0
Kluivert	0	0	0	0.602
Kocak	0	0	0	0.602
Latih	0	0	0	0
Lokal	0	0	0	0.602
Limpah	0	0	0.602	0
Main	0	0	0.602	0
Manfaat	0	0	0.602	0
Mantap	0	0.301	0.301	0
Prestasi	0	0	0	0.602
Stadion	0	0	0.602	0
Stok	0	0	0.602	0
Turun	0	0	0.602	0
Tunggu	0	0.602	0	0
Indonesia	0	0.602	0	0

Selanjutnya tahap ini setelah data sentimen dibersihkan dan terstruktur, kita melangkah maju menuju tahap klasifikasi. Dalam tahap ini, algoritma K-means Clustering bagaikan guru yang akan dilatih untuk membedakan hasil data positif, netral dan negatif.

a. Proses ini menggunakan empat dokumen pendek yang sudah direpresentasikan dalam vector TF-IDF (6 dimensi).

Tabel 13. Data Representase

Dokumen	Vektor TF-IDF (6 dimensi)
D1	[0.602, 0, 0, 0, 0.602, 0]
D2	[0, 0.301, 0, 0, 0, 0,]
D3	[0, 0.301, 1.204, 0, 0, 0]
D4	[0, 0, 0, 0.602, 0, 0]

b. Proses inisialisasi centroid dilakukan secara acak dengan memilih tiga dokumen awal sebagai centroid untuk setiap cluster sebanyak k 3 centroid awal:
   C1 = D1 [0.602, 0, 0, 0, 0.602, 0]
   C2 = D2 [0, 0.301, 0.602, 0, 0, 0]
   C3 = D4 [0, 0, 0, 0.602, 0, 0]

c. Mengukur jarak data ke setiap centroid, kemudian menentukan centroid terdekat untuk setiap data ke centroid dengan euclidean distance. Langkah pertama adalah menghitung jarak D1 ke centroid :
   D1 = [0.602, 0, 0, 0, 0.602, 0]

   D1 ke C1

$$\sqrt{(0.602-0.602)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0.602-0.602)^2+(0-0)^2}$$
$$=0$$

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

Setelah diketahui nilai D1 ke C1 sama = 0 , dan nilai nya sama persis dengan nilai D1, maka tidak perlu dilanjutkan untuk mencari nilai ke klaster C2 dan C3.

Selanjutnya menghitung jarak D2 ke centroid :

D2 = [0, 0.301, 0, 0, 0, 0,]

D2 ke C1

$$\sqrt{(0-0.602)^2+(0.301-0)^2+0+0+(0-0.602)^2}$$

$$=\sqrt{0.362+0.090+0+0+0.362}=\sqrt{2.263}\ =0.902$$

D2 ke C2

$$\sqrt{0^2+(0.301-0.301)^2+(0-0.602)^2}=\sqrt{0+0+0.362}=\sqrt{0.362}=0.602$$

D2 ke C3

$$\sqrt{0.301^2+0.602^2}=\sqrt{0.090+0.362}=\sqrt{0.362}=0.672$$

Setelah diketahui nilai D2 terhitung paling dekat ke C2 dengan nilai = 0.602 maka D3 ditetapkan di klaster 2

Langkah ini dilakukan sampai pada tahap menghitung D4 ke centroid sebagai berikut :

D4 ke C1

$$\sqrt{0.602^2+0+0+0.602^2+0.602^2}=\sqrt{0.362+0.362+0.362}=\sqrt{1.086}=1.042$$

D4 ke C2

$$\sqrt{0.301^2+0.602^2+0.602^2}=\sqrt{0.090+0.362+0.362}=\sqrt{1.086}=0.902$$

D4 ke C3

$$\sqrt{(0-0)^2+(0-0)^2+(0-0)^2+(0.602-0.602)^2+(0-0)^2\ +(0-0)^2}=0$$

Setelah diketahui nilai D4 terhitung paling dekat ke C3 dengan nilai = 0 maka D4 ditetapkan ke klaster 3

d. Mengelompokkan data berdasarkan seberapa dekatnya dengan centroid. Data yang memiliki nilai paling mendekati centroid akan diatributkan ke dalam cluster tersebut pada tabel 14.

Tabel 14. Data Pada Klaster

Dokumen	Jarak terdekat	klaster
D1	C1	1
D2	C2	2
D3	C2	2
D4	C3	3

e. Memperbarui nilai centroid dengan menghitung ulang jarak. Jika tidak ada perubahan pada pusat klaster, proses klasterisasi dianggap selesai. Namun, jika terdapat perubahan, langkah ke-4 harus diulangi hingga tidak ada lagi perubahan pada pusat klaster (centroid).

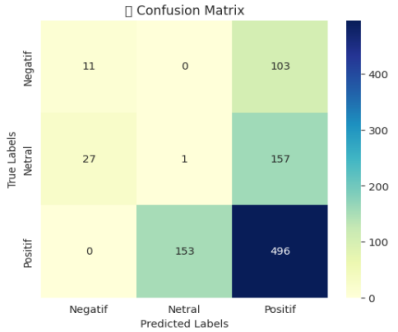
f. Iterasi dilakukan dari langkah 3 hingga langkah 5 hingga tidak ada perbedaan antara pusat centroid terakhir dan centroid sebelumnya. Iterasi terakhir menjadi acuan untuk menentukan pembagian cluster pada data dan proses sudah konvergen. Hasil klaster manual sesuai dengan dugaan sentimen yaitu negatif, netral, dan positif.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

Tabel 15. Hasil Uji Pada Penerapan K-Means Clustering

Dokumen	Tweet	Label Awal	Label Akhir
D1	buset latih kepala	Positif	Negatif
D2	mantap tunggu latih bawa indonesia	Positif	Netral
D3	main akademi mantap	Positif	Netral
D4	kluivert prestasi ajar latih lokal kocak	Positif	Positif

Dari hasil uji yang sudah dilakukkan dengan algoritma k means clustering maka didapat akurasi sebesar 25% karena sudah menebak 1 dari 4 dokumen yang diberikan. Setelah proses pelatihan dan pengujian, sistem menghasilkan label sentimen yang kemudian dibandingkan dengan label aktual untuk mengevaluasi kinerja model menggunakan metrik seperti akurasi, presisi, recall, dan f1-score. Karena distribusi data tidak seimbang dengan jumlah sentimen positif lebih dominan dibanding netral dan negatif evaluasi difokuskan pada sentimen positif. Hasil evaluasi ini divisualisasikan melalui confusion matrix untuk melihat sejauh mana algoritma K-Means Clustering mampu mengelompokkan sentimen dengan baik.



Gambar 4. Hasil Confusion Matrix

Pada gambar 4 hasil pengklasifikasian untuk mencari *accuracy*, *precision*, *recall* dan *f1-score* mengikuti rumus berikut :

Tabel 16. Penentuan Nilai TP, FP, TN Dan FP Kelas Positif

Label	Nilai
TP(*True positive*)	496
FP(*False positive*)	260
TN(*True Negative*)	97
FN(*False Negative*)	153

$$\text{Akurasi} = \frac{TP_{Total}}{Total\ Semua\ Data} = \frac{496+1+11}{948} = \frac{508}{948} = 0.5359 \text{ X } 100 = 53.59\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} = \frac{496}{496+260} = \frac{496}{756} = 0.656 \text{ X } 100 = 65.6\%$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} = \frac{496}{496+153} = \frac{496}{649} = 0.764 \text{ X } 100 = 76.4\ \%$$

$$\text{F1-Score} = \frac{\text{2 x Precision x Recall}}{\text{Precision+ Recall}} = \frac{2 \text{ x } 0.656 \text{ x } 0.764}{0.656+0.764} = \frac{1.002}{1.42} = 0.706 \text{ X } 100 = 70.6\%$$

Tabel 17. Penentuan Nilai TP, FP, TN Dan FP Kelas Negatif

Label	Nilai
TP(*True positive*)	11
FP(*False positive*)	27
TN(*True Negative*)	103
FN(*False Negative*)	865

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

$$\text{Akurasi} = \frac{TP_{Total}}{Total\ Semua\ Data} = \frac{496+1+11}{948} = \frac{508}{948} = 0.5359 \text{ X } 100 = 53.59\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} = \frac{11}{11+27} = \frac{11}{38} = 0.289 \text{ X } 100 = 28.9\%$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} = \frac{11}{11+103} = \frac{11}{114} = 0.096 \text{ X } 100 = 9.6\ \%$$

$$\text{F1-Score} = \frac{2 \text{ x Precision x Recall}}{\text{Precision+ Recall}} = \frac{2 \text{ x } 0.289 \text{ x } 0.096}{0.289 + 0.096} = \frac{0.0555}{0.385} = 0.144 \text{ X } 100 = 14.4\%$$

Tabel 18. Penentuan Nilai TP, FP, TN Dan FP Kelas Netral

Label	Nilai
TP(*True positive*)	1
FP(*False positive*)	153
TN(*True Negative*)	184
FN(*False Negative*)	668

$$\text{Akurasi} = \frac{TP_{Total}}{Total\ Semua\ Data} = \frac{496+1+11}{948} = \frac{508}{948} = 0.5359 \text{ X } 100 = 53.59\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} = \frac{1}{1+153} = \frac{1}{154} = 0.0065 \text{ X } 100 = 0{,}65\%$$

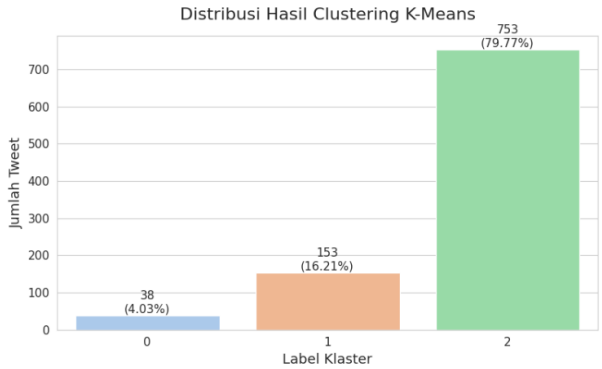
$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} = \frac{1}{1+184} = \frac{1}{185} = 0.0054 \text{ X } 100 = 0.54\ \%$$

$$\text{F1-Score} = \frac{2 \text{ x Precision x Recall}}{\text{Precision+ Recall}} = \frac{2 \text{ x } 0.0065 \text{ x } 0.0054}{0.0065 + 0.0054} = \frac{0.0000702}{0.0119} = 0.0059 \text{ X } 100 = 0.59\%$$

Secara keseluruhan nilai-nilai diatas dapat disajikan dalam classification report. Berikut disajikan classification report dari confusion matrix atas pengujian data uji dengan metode K-Means Clustering.

```
Akurasi: 53.59 %

Classification Report:
              precision    recall  f1-score   support

     Negatif       0.29      0.10      0.14       114
      Netral       0.01      0.01      0.01       185
     Positif       0.66      0.76      0.71       649

    accuracy                           0.54       948
   macro avg       0.32      0.29      0.29       948
weighted avg       0.49      0.54      0.50       948

Adjusted Rand Index (ARI): -0.025
Silhouette Score: 0.104
```

Gambar 5. Hasil Classification Report



Gambar 6. Hasil K-means Clustering

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

4. KESIMPULAN

Berdasarkan hasil penelitian, analisis sentimen terhadap 947 data komentar dari media sosial X dilakukan menggunakan metode K-Means Clustering. Pelabelan sentimen dilakukan dengan pendekatan lexicon-based, yang membagi data ke dalam tiga kategori: positif, netral, dan negatif berdasarkan skor polaritas teks. Hasil evaluasi menunjukkan akurasi model sebesar 53,59%, dengan performa cukup baik pada kelas positif (precision 0.66, recall 0.76, f1-score 0.71), namun sangat rendah pada kelas netral dan negatif. Temuan ini menunjukkan bahwa K-Means kurang efektif dalam mengelompokkan data dengan distribusi tidak seimbang, sehingga lebih cocok digunakan sebagai baseline awal dalam analisis sentimen teks.

DAFTAR PUSTAKA

[1] M. A. Akbar, F. Fatimah, and J. Jaenudin, "Penerapan Data Mining Untuk Pengelompokan Posisi Pemain Sepak Bola Menggunakan Algoritma K-Means Clustering," 2019. doi:10.31849/digitalzone.v14i2.17106

[2] N. Fitriyah, B. Warsito, D. Asih, and I. Maruddani, "ANALISIS SENTIMEN GOJEK PADA MEDIA SOSIAL TWITTER DENGAN KLASIFIKASI SUPPORT VECTOR MACHINE (SVM)," *J. GAUSSIAN*, vol. 9, no. 3, pp. 376–390, 2020. doi:/10.14710/j.gauss.9.3.376-390

[3] M. Fajar, N. Rahaningsih, and R. Danar Dana, "Analisis Pola Penjualan Obat Di Apotek an-Naafi Menggunakan Metode K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 486–492, 2024, doi: 10.36040/jati.v8i1.8395.

[4] M. Refa, T. Ramdhani, M. Ridho, P. Budika, M. V. Santoso, and N. Zahry, "Analisis Sentimen Terhadap Komentar Negatif (Hate Speech) Di Twitter Dengan Algoritma K-Means Clustering Menggunakan Rapidminer," vol. 1, no. 1, pp. 57–61, 2025. doi:/10.31294/ji.v7i1.9601

[5] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.

[6] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat," *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, vol. 6, no. 2, p. 48, 2021, doi: 10.51557/pt_jiit.v6i2.659.

[7] M. Sholeh and K. Aeni, "Perbandingan Evaluasi Metode Davies Bouldin, Elbow dan Silhouette pada Model Clustering dengan Menggunakan Algoritma K-Means," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 8, no. 1, p. 56, 2023, doi: 10.30998/string.v8i1.16388.

[8] M. A. Rheza and F. Metandi, "Implementasi Metode K-Means Clustering Untuk Penentuan Jenis Komentar Pada Tweet Pssi," *Just TI (Jurnal Sains Terap. Teknol. Informasi)*, vol. 12, no. 2, p. 73, 2020, doi: 10.46964/justti.v12i2.363.

[9] T. Agustian, E. Fresia Nandela, S. A. Sinay, M. Habibi, and F. Unjaya, "Pemetaan Opini Publik Menggunakan Data Mining: Studi Kasus Naturalisasi Pemain Sepak Bola dengan K-Means dan Naive Bayes Classifier," 2024. doi:/10.19184/e-sospol.v12i1.53714

[10] A. Chusyairi and P. Ramadar Noor Saputra, "Pengelompokan Data Puskesmas Banyuwangi Dalam Pemberian Imunisasi Menggunakan Metode K-Means Clustering," *Telematika*, vol. 12, no. 2, pp. 139–148, 2019, doi: 10.35671/telematika.v12i2.848.

[11] M. Ulfah and A. Sri Irtwaty, "Penerapan Data Mining Clustering Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan," *Fidel. J. Tek. Elektro*, vol. 4, no. 3, pp. 62–68, 2022, doi: 10.52005/fidelity.v4i3.126.

[12] H. A. Ulvi and M. Ikhsan, "Comparison of K-Means and K-Medoids Clustering Algorithms for Export and Import Grouping of Goods in Indonesia," *J. dan Penelit. Tek.*

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503

Inform., vol. 8, no. 3, pp. 1641–1655, 2024, [Online]. Available: https://doi.org/10.33395/sinkron.v8i3.13815

[13] H. P. Kurniawan and L. Farhatuaini, "Identifikasi Pola Kepuasan Mahasiswa Terhadap Proses Pembelajaran Menggunakan Algoritma K-Means Clustering.," *J. Inform. J. Pengemb. IT*, vol. 9, no. 2, pp. 164–172, 2024, doi: 10.30591/jpit.v9i2.6740.

[14] S. Bila Rahmania Sharafi, "Identifikasi Pola Diskusi Publik mengenai Pemindahan Ibu Kota Negara Menggunakan Analisis TF-IDF dan K-Means Clustering," *Semin. Nas. Sist. Inf.*, vol. 08, pp. 4654–4663, 2024, [Online]. Available: https://jurnalfti.unmer.ac.id/index.php/senasif/article/view/557. doi:/10.30599/c0zqdw84

[15] D. Sabrina, A. D. Sabilla, and N. Azizah, "KOMBINASI VADER LEXICON DAN SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI SENTIMEN KOMENTAR APLIKASI BLU BCA," vol. 6, no. 1, pp. 22–33, 2025. doi:/10.23887/insert.v6i1.86240

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13503