

Analisis Pemain Terbaik Sepak Bola dengan menggunakan Algoritma K-Means

Faviola Proba Wardhana^{1,*}, Sri Winarno¹

¹ Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Indonesia

* Correspondence: 111202013197@mhs.dinus.ac.id

Copyright: © 2024 by the authors

Received: 23 Juli 2024 | Revised: 27 Juli 2024 | Accepted: 14 Oktober 2024 | Published: 19 Desember 2024

Abstrak

Pemilihan pemain dalam olahraga sepak bola merupakan salah satu hal yang penting untuk penyusunan strategi dalam pertandingan. Hal ini menjadi sistem pendukung keputusan yang dapat digunakan untuk melakukan pemilihan *line-up* pemain. Data pada penelitian ini diperoleh dari situs resmi Liga 1 Indonesia musim 2023 sepak bola yang terpercaya. Penelitian ini bertujuan untuk menganalisis pemain terbaik sepak bola dengan menggunakan algoritma *k-means* berdasarkan performa statistik selama Liga 1 Indonesia musim 2023. Pengumpulan data dari setiap pemain, diambil jumlah presentasi sebagai pemain utama (*starting line up*). Kami menggunakan Algoritma *k-means*. *K-means* dapat digunakan untuk mengidentifikasi pola dan mengelompokkan pemain berdasarkan instrumen statistik pertandingan seperti jumlah gol, assist, dan statistik fisik lainnya disetiap posisi pemain. Data yang diperoleh sebanyak 197 pemain yang berkompetisi pada Liga 1 2023. Hasil temuan kami adalah terdapat 62 pemain dalam *cluster* 1 dari total 197 pemain berdasarkan analisis yang telah dilakukan. Pemain-pemain ini memiliki statistik terbaik dan bisa menjadi pilihan bagi staf kepelatihan klub Liga 1 untuk direkrut atau dibeli guna memperkuat tim di musim berikutnya. Penelitian kami menunjukkan bahwa para pemain dalam kluster ini memiliki performa yang menonjol, sehingga dapat membantu staf pelatih dalam mengidentifikasi tipe pemain seperti "penyerang efisien" atau "bek tangguh". Dengan demikian, penelitian ini dapat mendukung pelatih atau manajer dalam memilih pemain yang paling sesuai dengan kebutuhan tim untuk musim mendatang.

Kata kunci: algoritma k-means; pemain sepak bola; performa

Abstract

The selection of players in soccer is crucial for developing strategies in matches. It serves as a decision support system that can be used to choose the starting line-up. The data for this research was obtained from the official and reliable website of Liga 1 Indonesia for the 2023 season. This study aims to analyze the top soccer players using the K-means algorithm based on their statistical performance throughout the 2023 Liga 1 Indonesia season. Data collection for each player included their percentage of appearances in the starting line-up. We used the K-means algorithm, which helps identify patterns and cluster players based on statistical metrics from the matches, such as the number of goals, assists, and other physical statistics across various player positions. The data comprised 197 players competing in Liga 1 2023. Our findings reveal that 62 players belong to Cluster 1 out of the total 197 analyzed. These players exhibited the best statistics and could be potential options for Liga 1 coaching staff to recruit or sign in order to strengthen their teams for the next season. Our research indicates that the players in this cluster demonstrated outstanding performance, helping coaches identify categories such as "efficient strikers" or "strong defenders." Therefore, this study can assist coaches or managers in selecting the most suitable players to meet the team's needs for the upcoming season.

Keywords: *k-means algorithm; football player; performance*



PENDAHULUAN

Sepak bola merupakan cabang olahraga yang menggunakan bola yang umumnya terbuat dari kulit dan dimainkan oleh dua tim masing-masing beranggotakan 11 (sebelas) orang. Permainan sepak bola sering sekali dimainkan oleh masyarakat untuk kebugaran, rekreasi dan prestasi, sehingga tidak heran mengapa sepak bola menjadi cabang olahraga yang sangat populer dan digemari oleh seluruh lapisan masyarakat (Rizal et al., 2019).

Salah satu tantangan besar yang dihadapi pelatih dan manajemen klub adalah menentukan pemain baru untuk musim mendatang. Biasanya, mereka bergantung pada data statistik mentah dan informasi subjektif yang tidak selalu mencerminkan performa pemain secara menyeluruh. Pada umumnya ketika pemain sepakbola mempunyai statistik performance yang baik maka pemain tersebut akan menjadi acuan untuk memainkan pemain tersebut. (Kaukab, et al., 2022). Statistik dasar seperti jumlah gol, *assist*, dan penyelamatan sering kali tidak memberikan gambaran lengkap tentang kualitas dan peran seorang pemain dalam tim, menyebabkan kesulitan dalam menilai dan mengelompokkan pemain dengan efektif. Hal ini mengakibatkan kesulitan dalam menilai dan mengelompokkan pemain secara efektif dan menjelaskan faktor-faktor apa saja yang memengaruhi peningkatan atau penurunan rating mereka dalam pertandingan-pertandingan tersebut (Rahman, et al., 2024).

Penelitian ini berfokus pada pengelompokan pemain sepak bola berdasarkan performa statistik mereka, yang selama ini belum optimal dalam berbagai penelitian sebelumnya. Banyak studi yang hanya menganalisis performa pemain secara umum tanpa mengelompokkan mereka secara rinci ke dalam tipe atau posisi tertentu. Hal ini menyebabkan kesulitan dalam memahami karakteristik spesifik tiap pemain, yang berdampak pada pengambilan keputusan kurang akurat dalam rekrutmen dan pengelolaan tim.

Penelitian ini menawarkan sebuah pendekatan inovatif dengan mengimplementasikan algoritma *k-means* untuk mengelompokkan pemain berdasarkan indikator statistik performa yang relevan, seperti jumlah gol, assist, intersepsi, dan akurasi umpan. Proses dimulai dengan normalisasi data untuk memastikan keseragaman skala antar variabel, yang kemudian diikuti oleh penerapan machine learning. *Machine learning* adalah cabang dari kecerdasan buatan yang memungkinkan sistem untuk belajar dan membuat keputusan atau prediksi tanpa diprogram secara eksplisit (Nugraha et al., 2023; Mulyadi & Purnomo, 2023; Sulastri et al., 2021), dan salah satunya adalah algoritma K-Means.

Algoritma *k-means* merupakan salah satu teknik *partition-based clustering* yang bertujuan untuk mengelompokkan data ke dalam kluster berdasarkan jarak dari centroid (Gosari & Rismayani, 2023; Qirom et al., 2024; Salam et al., 2020), di mana setiap pemain yang memiliki kesamaan performa akan dikelompokkan bersama (Andiani et al., 2022; Hafiz et al., 2024; Harahap et al., 2022; Martanto & Hayati, 2024). Algoritma ini telah terbukti efisien dalam memisahkan data berdasarkan pola, sehingga memudahkan dalam mengidentifikasi performa pemain yang sejenis (Ashari et al., 2022; Wadanur & Sari, 2022). Namun, untuk memastikan hasil yang akurat, normalisasi data menjadi langkah penting. Normalisasi ini dilakukan untuk menghindari bias yang mungkin terjadi akibat perbedaan skala antar variabel, seperti gol dan *assist* (Augusto et al., 2022). Dengan normalisasi, setiap indikator performa pemain akan berkontribusi secara proporsional dalam analisis, memastikan bahwa tidak ada variabel yang mendominasi hasil *clustering*, sehingga hasil akhir lebih seimbang dan representatif.

Penelitian sebelumnya yang dilakukan oleh García, et al. (2020), menunjukkan bahwa penggunaan algoritma *k-means* berhasil mengidentifikasi pemain kiper dengan jumlah gol kemasukan yang rendah, sementara (García-Aliaga, et al., 2021), menemukan bahwa bek dalam *Cluster* yang sama memiliki tingkat intersepsi dan akurasi umpan yang tinggi. Dengan demikian, pendekatan ini tidak hanya memperjelas pemahaman tentang performa pemain, tetapi juga membantu dalam pengambilan keputusan yang lebih tepat dalam proses rekrutmen.

Perbedaan utama antara penelitian ini dengan penelitian sebelumnya adalah terletak pada pendekatan dan metodologi yang digunakan dalam menganalisis performa pemain sepak bola. Penelitian kami mengadopsi algoritma *k-means* untuk melakukan pengelompokan pemain berdasarkan berbagai dimensi performa, seperti gol, *assist*, dan intersepsi. Berbeda dari penelitian sebelumnya yang cenderung menggunakan data statistik mentah atau teknik *Clustering* yang kurang mendalam, penelitian ini menyempurnakan metode tersebut dengan menekankan pentingnya normalisasi data (Nurzahputra et al., 2017; Shelly et al., 2020). Normalisasi dilakukan agar setiap indikator performa memiliki kontribusi yang seimbang dalam proses *clustering*. Kesenjangan yang diisi oleh penelitian ini adalah kurangnya penerapan normalisasi secara sistematis dalam studi-studi sebelumnya. Akibatnya, hasil pengelompokan seringkali tidak proporsional karena beberapa variabel, seperti gol atau *assist*, mendominasi hasil. Dengan melakukan normalisasi, penelitian ini memastikan bahwa semua variabel, baik itu gol, *assist*, atau intersepsi diperlakukan secara setara, sehingga memberikan hasil yang lebih akurat dan representatif mengenai performa pemain.

Inovasi ini bertujuan untuk memberikan wawasan yang lebih mendalam dan terperinci tentang kualitas pemain, sekaligus memperbaiki keterbatasan yang ada dalam penelitian sebelumnya. Melalui penerapan algoritma *k-means* yang disertai normalisasi data yang tepat, penelitian ini mampu menghasilkan kluster yang lebih akurat dan representatif. Hal ini memungkinkan pemahaman yang lebih mendalam tentang kemampuan dan potensi pemain, terutama dalam konteks rekrutmen strategis dan pengambilan keputusan yang lebih tepat.

Penelitian ini bertujuan untuk melakukan analisis pemain terbaik sepak bola dengan menggunakan algoritma *k-means*. Hasil analisis tidak hanya akan menawarkan wawasan yang lebih rinci terkait statistik performa, tetapi juga akan membantu pelatih dan manajemen klub dalam merumuskan strategi rekrutmen yang lebih efektif dan berbasis informasi untuk musim 2024. Dengan pendekatan ini, penelitian ini berpotensi mendukung pengambilan keputusan yang lebih baik, meningkatkan kualitas tim, dan memastikan bahwa klub dapat merekrut pemain yang sesuai dengan kebutuhan serta visi mereka. Selain itu, penelitian ini diharapkan dapat mengisi kekurangan dalam studi sebelumnya dengan menghadirkan metodologi yang lebih canggih, sehingga memberikan kontribusi signifikan terhadap pengembangan analisis performa dalam sepak bola.

METODE

Data penelitian ini diperoleh dari situs statistik sepakbola musim Liga 1 2023 pada aplikasi FotHub. Pengumpulan data setiap pemain, akan diambil jumlah presentasi sebagai pemain utama (*starting line up*), jumlah pertandingan satu musim dan menit bermain. Setiap posisi pemain memiliki *clustering* yang berbeda, sebagai berikut. Kiper: Persentase penyelamatan, jumlah gol yang diterima, jumlah clean sheet (tidak kebobolan), dan distribusi umpan. Bek: Jumlah intersep, tekel sukses, persentase duel yang dimenangkan, dan persentase keberhasilan umpan. Gelandang: Jumlah assist, umpan kunci, jumlah tendangan, persentase keberhasilan umpan, dan jumlah tekel. Striker: Jumlah gol, jumlah tendangan ke gawang, jumlah assist, dan efisiensi finishing.

Pengumpulan data dari setiap pemain akan diambil jumlah presentasi sebagai pemain utama (*starting line up*). Posisi pemain yang digunakan dalam penelitian ini meliputi penjaga gawang (kiper), pemain belakang (bek), pemain tengah (gelandang), dan pemain depan (*striker*). Berdasarkan analisis, atribut yang digunakan untuk kiper meliputi banyaknya gol kemasukan, penyelamatan, dan clean sheet. Pada pemain belakang, atribut yang digunakan adalah intercept, jumlah tekel, dan persentase keberhasilan umpan. Untuk pemain tengah, atribut yang digunakan meliputi jumlah assist, jumlah tendangan, jumlah tekel, dan persentase keberhasilan umpan. Untuk pemain depan, atribut yang digunakan meliputi jumlah gol, jumlah tendangan ke gawang, assist, dan efisiensi finishing. Atribut umur, klub asal, dan jumlah

pertandingan tidak digunakan dalam perhitungan *k-means*. Adapun attribute tranformation data ini dapat dilihat pada tabel 1.

Tabel 1. *Attribute transformation*

No	Posisi Pemain	Atribut awal	Atribut baru
1.	Kiper	Gol Kemasukan	GK
		Penyelamatan	SF
		<i>Clean Sheet</i>	CS
2.	Bek	<i>Intercept</i>	IC
		Jumlah Tekel	JT
		Keberhasilan Umpan	KU
3.	Gelandang	<i>Assist</i>	AS
		Jumlah Tendangan	JS
		Jumlah Tekel	JT
		Keberhasilan Umpan	KU
4.	Stiker	Jumlah Gol	GL
		Tendangan Ke Gawang	ST
		<i>Assist</i>	AS
		Efisiensi Finishing	EF

Langkah awal yang dilakukan adalah menentukan jumlah *Cluster* yaitu sebanyak 3 *Cluster*. Langkah selanjutnya menentukan pusat awal *Cluster (centroid)*. Ini dilakukan dengan mengambil data dari hasil statistik pertandingan yang telah dilakukan oleh para pemain pada setiap posisi. Penentuan *centroid* ini dilakukan dengan menggunakan data statistik dari penampilan pemain selama pertandingan yang telah berlangsung.

Setelah *centroid* telah ditentukan dan data didapatkan tahapan selanjutnya data akan dilakukan *preprocessing* atau pembersihan data, dilakukan deteksi dan perbaikan atau penghapusan data yang tidak relevan, *missing value*, dan data redundan dengan cara mengatur kembali data yang ada pada tabel. Pembersihan data dilakukan secara manual dengan memeriksa data satu per satu dan menghapus data yang tidak relevan, *missing value*, dan data redundan. Selanjutnya, pada tahap transformasi, dilakukan perubahan bentuk atribut dari data yang didapat supaya sesuai dengan proses data mining.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i + y_i)^2} \tag{1}$$

Pada penelitian ini perhitungan jarak yang digunakan adalah *euclidean distance* yang merupakan perhitungan jarak dari 2 titik dalam *euclidean space*. Setelah didapatkan jarak maka akan di ambil jarak yang terpendek. Adapun tahap analisa peforma pemain sepak bola menggunakan algoritma *k-means*. Persamaan (1) digunakan untuk menghitung jarak antara data pemain (titik data) dan pusat *cluster (centroid)*. Tujuannya adalah untuk menentukan seberapa dekat seorang pemain dengan setiap centroid dari *cluster* yang berbeda. Fungsi persamaan dalam *k-means* sangat penting karena menentukan di *cluster* mana pemain tersebut akan dikelompokkan. Pemain akan ditempatkan ke dalam *cluster* yang memiliki jarak terdekat dari *centroid*. Setelah semua pemain ditempatkan, algoritma *k-means* akan memperbarui posisi *centroid* berdasarkan rata-rata pemain yang berada dalam *cluster* tersebut, lalu iterasi ini berlanjut hingga *cluster* stabil.

HASIL DAN PEMBAHASAN

Hasil

Analisis menggunakan algoritma *k-means* bertujuan untuk mengelompokkan data ke dalam beberapa *cluster* berdasarkan karakteristik yang memiliki kesamaan. Pada kasus ini, pemain dari berbagai posisi dikelompokkan berdasarkan atribut-atribut yang relevan, seperti usia, jumlah penyelamatan, intercept, assist, gol, dan efisiensi. Setiap *cluster* mewakili segmen pemain dengan karakteristik yang homogen dalam *cluster* tersebut, tetapi *heterogen* antar *cluster*. Setelah *centroid* awal ditentukan, hitung jarak antara setiap data pemain dengan pusat *cluster* (*centroid*).

Tabel 2. Hasil interpretasi algoritma *k-means*

No	Posisi Pemain	Identifikasi	Cluster 1	Cluster 2	Cluster 3
1.	Kiper	Jumlah Anggota	6	8	4
		Rata-rata usia	25,84	28,38	26,5
		Gol Kemasukan	17,84	38,63	29,75
		Penyelamatan	2	6	9,25
		<i>Clean Sheet</i>	2	6	9,25
2.	Bek	Jumlah Anggota	21	27	23
		Rata-rata usia	27,76	29,59	28,57
		<i>Intercept</i>	31,48	132,70	75,26
		Jumlah Tekel	18,95	47,33	44,83
		Keberhasilan Umpan	81,30	84,87	80,28
3.	Gelandang	Jumlah Anggota	20	26	8
		Rata-rata usia	29,35	28,46	28,13
		<i>Assist</i>	5,45	1,38	2,5
		Jumlah Tendangan	44,8	10,50	16,88
		Jumlah Tekel	28,35	23,77	74,75
4.	Stiker	Keberhasilan Umpan	76,75	83,04	85,79
		Jumlah Anggota	14	21	19
		Rata-rata usia	30,29	27,05	28,95
		Rata-rata Jumlah Gol	15,71	1,76	5,95
		Tendangan Ke Gawang	36,86	7,24	20,16
		<i>Assist</i>	5,21	1,90	2,89

Pada posisi kiper, *k-means* membagi pemain menjadi tiga *cluster* berdasarkan usia, jumlah gol kemasukan, jumlah penyelamatan, dan jumlah clean sheet. *Cluster 1* memiliki pemain dengan usia rata-rata lebih muda dan catatan penyelamatan yang lebih sedikit, sedangkan *cluster 2* dan 3 menunjukkan peningkatan usia serta kemampuan penyelamatan yang lebih baik. Pada pemain belakang (Bek), analisis *k-means* mengelompokkan mereka berdasarkan intercept, tekel, dan persentase keberhasilan umpan. *Cluster 2*, misalnya, berisi pemain dengan jumlah *intercept* dan tekel yang jauh lebih tinggi dibanding *cluster 1*, mengindikasikan peran mereka sebagai bek yang lebih defensif. Sedangkan *cluster 3* memiliki keseimbangan antara intercept, tekel, dan keberhasilan umpan. Pada pemain tengah (Gelandang), pengelompokan berdasarkan assist, jumlah tendangan, tekel, dan keberhasilan umpan menunjukkan peran yang berbeda. *cluster 1* cenderung berisi pemain dengan kontribusi assist yang lebih besar, sementara *cluster 2* lebih fokus pada distribusi bola dengan keberhasilan umpan yang lebih tinggi. Terakhir, pada pemain depan (*Striker*), *k-means* mengelompokkan mereka berdasarkan jumlah gol, tendangan ke gawang (*shot on target*), assist, dan efisiensi finishing. *Cluster 1* menampilkan striker dengan performa gol dan efisiensi

finishing yang tinggi, sementara *cluster 2* memiliki striker dengan kontribusi gol dan shot on target yang lebih rendah, menunjukkan tingkat produktivitas yang berbeda di setiap *cluster*. Analisis ini memungkinkan pelatih atau analis memahami pola-pola tersembunyi dalam data pemain dan menggunakannya untuk pengambilan keputusan strategis.

Pada tabel 2 menunjukkan bahwa pemain dalam *cluster 1* di setiap posisi memiliki performa yang paling menonjol dibandingkan *cluster* lainnya. Pada posisi kiper, meskipun jumlah *clean sheets* mereka rendah, mereka memiliki kemampuan pertahanan yang baik dengan gol kemasukan paling sedikit. Di sisi lain, kiper dalam *cluster 3* unggul dalam jumlah penyelamatan dan *clean sheets*, tetapi menghadapi lebih banyak kebobolan, yang menunjukkan bahwa mereka mungkin lebih sering diuji.

Pada posisi bek, pemain di *cluster 1* memiliki persentase keberhasilan umpan yang solid, meskipun jumlah *intercept* mereka lebih rendah. Sementara itu, *cluster 2* menunjukkan performa yang sangat agresif dalam bertahan dengan *intercept* dan tekel yang jauh lebih tinggi, menjadikannya bek yang lebih kuat dalam pertahanan. Untuk gelandang, *cluster 1* unggul dalam kontribusi serangan, baik dalam hal jumlah tendangan maupun assist. *Cluster 2*, meski lebih rendah dalam assist, menunjukkan distribusi bola yang sangat baik dengan keberhasilan umpan tertinggi, menandakan mereka lebih berfokus pada peran pengatur permainan. *Cluster 3* menonjol dalam aspek bertahan di lini tengah, dengan jumlah tekel yang jauh lebih banyak. Pada posisi striker, *cluster 1* jelas unggul dalam produktivitas gol, *assist*, dan efisiensi finishing, menjadikannya opsi terbaik dalam serangan. Sementara itu, *cluster 3* menunjukkan performa sedang, dan *cluster 2* kurang efektif dalam mencetak gol. Secara keseluruhan, pemain di *cluster 1* adalah yang paling komprehensif dalam hal performa, dan karena itu, mereka dapat diprioritaskan untuk direkrut pada Liga 1 musim 2024, mengingat kontribusi mereka yang signifikan pada musim 2023.

Pembahasan

Hasil analisis menggunakan algoritma *k-means* dalam penelitian ini berhasil mengelompokkan pemain sepak bola berdasarkan performa mereka selama musim Liga 1 2023. Proses pengelompokan ini didasarkan pada berbagai indikator statistik seperti gol, assist, intersepsi, dan statistik fisik lainnya, yang telah diolah dan dinormalisasi. Melalui pendekatan ini, penelitian bertujuan untuk mengidentifikasi pola performa pemain dengan lebih akurat, memberikan wawasan yang mendalam mengenai kinerja mereka, dan memfasilitasi pengambilan keputusan yang lebih strategis oleh tim pelatih dalam menyusun strategi dan rekrutmen pemain.

Pada pemain dengan posisi kiper terdapat 6 data pemain pada *cluster 1* dengan usia rata-rata 25,84 tahun. Pada *cluster 2* terdapat 8 data pemain dengan usia rata-rata 28,38 tahun. Pada *cluster 3* dengan posisi pemain sebagai kiper terdapat 4 data pemain dengan usia rata-rata 26,5 tahun. Informasi yang menonjol pada *cluster 1* yaitu banyaknya gol kemasukan yang lebih sedikit dibandingkan dengan *cluster 2* dan *cluster 3*. Informasi yang menonjol pada *cluster 3* yaitu penyelamatan dan *clean sheet* yang lebih tinggi dibandingkan dengan *cluster 1* dan *cluster 2*.

Pada pemain dengan posisi sebagai pemain belakang terdapat 21 data pemain dengan usia rata-rata 27,76 tahun. Pada *cluster 2* terdapat 27 data pemain dengan usia rata-rata 29,59 tahun. Informasi yang menonjol pada *cluster 2* yaitu banyaknya persentase keberhasilan umpan yang lebih tinggi dibandingkan dengan *cluster 1* dan *cluster 3*. Kemudian pada *cluster 3* dengan posisi pemain sebagai bek terdapat 23 data pemain dengan usia rata-rata 28,57 tahun.

Pada pemain tengah (Gelandang) diketahui bahwa *cluster 1* terdapat 20 data pemain dengan usia rata-rata 29,35 tahun. Pada *cluster 2* terdapat 26 data pemain dengan usia rata-rata 28,46 tahun. Kemudian pada *cluster 3* dengan posisi pemain sebagai gelandang terdapat 8 data pemain dengan usia rata-rata 28,13 tahun. Informasi yang menonjol pada *cluster 1* yaitu

banyaknya rata-rata jumlah tendangan dan *assist* yang lebih tinggi dibandingkan dengan *cluster 2* dan *cluster 3*.

Selanjutnya, pada pemain depan (*Stiker*) diketahui bahwa *cluster 1* terdapat 14 data pemain dengan usia rata-rata 30,29 tahun. Pada *cluster 2* terdapat 21 data pemain dengan usia rata-rata 27,05 tahun. Kemudian pada *cluster 3* dengan posisi pemain sebagai striker terdapat 19 data pemain dengan usia rata-rata 28,95 tahun. Informasi yang menonjol pada *cluster 1* yaitu banyaknya jumlah gol, jumlah tendangan, assist dan efisiensi *finishing* yang lebih tinggi dibandingkan dengan *cluster 2* dan *cluster 3*.

Algoritma *k-means* dipilih karena kesederhanaannya dan kemampuannya untuk bekerja dengan cepat dan efisien pada dataset yang relatif kecil hingga sedang. Algoritma ini memberikan hasil yang mudah diinterpretasikan melalui penghitungan rata-rata karakteristik setiap *cluster*. Dalam kasus ini, data performa pemain yang terstruktur cocok untuk diolah dengan *k-means*, karena algoritma ini berfokus pada pengelompokan data yang memaksimalkan kesamaan dalam *cluster* dan perbedaan antar *cluster*. Teori *partition-based cluster* mendukung penggunaan *k-means* untuk membagi data berdasarkan kedekatan dengan *centroid cluster*, yang menjadikannya tepat untuk memahami karakteristik pemain berdasarkan performa mereka.

Temuan sebelumnya yang dilakukan oleh Huang & Wang (2019) serta García & Sanchis (2020), dalam kajian sebelumnya, *cluster 1* diidentifikasi sebagai kelompok yang terdiri dari pemain dengan statistik unggulan dalam hal gol, assist, dan keterlibatan defensif. Keunggulan penelitian ini dibandingkan dengan penelitian sebelumnya terletak pada penerapan normalisasi data yang lebih baik dan fokus pada pengelompokan berdasarkan posisi spesifik. Normalisasi data menjadi penting untuk memastikan setiap indikator performa memberikan kontribusi yang seimbang dalam proses *clustering*, sehingga hasil yang diperoleh lebih representatif. Penelitian sebelumnya cenderung mengabaikan pentingnya normalisasi ini, yang dapat menyebabkan ketidakseimbangan dalam hasil *clustering*. Sementara itu, temuan kami menyempurnakan kekurangan yang ada pada penelitian sebelumnya dengan lebih spesifik dalam mengidentifikasi kemampuan pemain di berbagai posisi. Kiper dalam *cluster 1* tercatat memiliki jumlah gol kemasukan yang lebih sedikit, sementara bek menunjukkan keunggulan dalam intersepsi dan akurasi umpan. Gelandang dan striker, di sisi lain, memperlihatkan performa superior dalam aspek *assist* dan gol. Oleh karena itu, hasil ini menegaskan keandalan metode *k-means* dalam mengelompokkan pemain berdasarkan kinerja yang terukur.

Selain itu, dampak penelitian ini pada bidang sepak bola, khususnya dalam konteks pengambilan keputusan rekrutmen sangatlah signifikan. Pengelompokan pemain yang lebih terperinci memberikan pelatih dan manajemen klub alat yang lebih komprehensif dan andal untuk menilai serta memilih pemain sesuai kebutuhan spesifik di setiap posisi. Dengan demikian, perencanaan strategis menjadi lebih akurat dan berbasis data, memungkinkan klub untuk memaksimalkan potensi tim melalui perekrutan pemain yang memiliki kinerja terbaik di posisi yang sesuai. yang relevan.

SIMPULAN

Hasil analisis algoritma *k-means* menunjukkan bahwa pemain dalam *cluster 1* di setiap posisi memiliki performa paling menonjol dibandingkan dengan *cluster* lainnya. Kiper di *cluster 1* memiliki gol kemasukan paling sedikit, bek menunjukkan persentase umpan yang efektif, gelandang unggul dalam tendangan dan *assist*, serta striker memiliki produktivitas gol dan efisiensi *finishing* tertinggi. Temuan ini menegaskan bahwa *cluster 1*, dengan keseimbangan antara usia, pengalaman, dan performa, adalah kandidat utama untuk rekrutmen di Liga 1 musim 2024. Dengan pemilihan ini, tim dapat mengoptimalkan potensi pemain berdasarkan data performa terbaik dari musim sebelumnya. Algoritma *k-means*, dengan

kemampuannya membagi data secara efektif, tetap menjadi alat yang tepat untuk identifikasi pemain, memudahkan pelatih dalam membuat keputusan rekrutmen yang strategis.

REFERENSI

- Andiani, D., Dwi, S., Septiani, R., & Riana, A. (2022). Analisis Teknik non-Hierarki untuk Pengelompokan Kabupaten/Kota di Provinsi Jawa Barat Berdasarkan Indikator Kesejahteraan Rakyat 2020. *Jurnal Riset Matematika Dan Sains Terapan*, 21(1), 21–28.
- Ashari, I. A., Negara, I. S. M., & Sumantri, R. B. B. (2022). Evaluasi Pembayaran Keuangan Siswa berdasarkan Penghasilan Wali Siswa menggunakan Metode Clustering K-Means. *Edumatic: Jurnal Pendidikan Informatika*, 6(2), 324-333. <https://doi.org/10.29408/edumatic.v6i2.6395>
- Augusto, D., Brito, J., Aquino, R., Paulucio, D., Figueiredo, P., Bedo, B. L. S., ... & Vasconcellos, F. (2022). Contextual variables affect peak running performance in elite soccer players: A brief report. *Frontiers in Sports and Active Living*, 4, 966146. <https://doi.org/10.3389/fspor.2022.966146>
- Fitri, E. N., Winarno, S., Budiman, F., Rohmani, A., Zeniarja, J., & Sugiarto, E. (2023). Decision Tree Simplification Through Feature Selection Approach in Selecting Fish Feed Sellers. *Jurnal Teknik Informatika (Jutif)*, 4(2), 301-309. <https://doi.org/10.52436/1.jutif.2023.4.2.747>
- García, J., & Sanchis, E. (2020). Applying Machine Learning Algorithms to Assess Player Performance in Soccer. *International Journal of Sports Science and Coaching*, 15(3), 459-471.
- García-Aliaga, A., Marquina, M., Coteron, J., Rodríguez-González, A., & Luengo-Sanchez, S. (2021). In-game behaviour analysis of football players using machine learning techniques based on player statistics. *International Journal of Sports Science & Coaching*, 16(1), 148-157. <https://doi.org/10.1177/1747954120959762>
- Gosari, N. C., & Rismayani, R. (2023). Penerapan Data Mining Dalam Mengelompokkan Kunjungan Wisatawan Mancanegara Di Prov. Sulawesi Selatan Dengan K-Means Dan SVM. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(3), 174-180. <https://doi.org/10.30591/jpit.v8i3.4554>
- Hafiz, M., Rahman, T., & Susanto, D. (2024). Klasterisasi penyakit menular di Indonesia menggunakan metode K-Means Clustering. *Journal of Computer*, 4(1), 50-57. <https://doi.org/10.33330/j-com.v4i1.3033>
- Harahap, L. M., Siregar, T., & Gultom, F. (2022). Klastering sayuran unggulan menggunakan algoritma K-Means. *Jurnal Teknik Informatika dan Sistem Informasi*, 8(3), 567-579. <https://doi.org/10.28932/jutisi.v8i3.5277>
- Kaukab, M. E. (2022). Football Player Market Value: Apakah Usia Pemain Berperan Dalam Penentuan Harga Pasar?. *Jurnal Penelitian Dan Pengabdian Kepada Masyarakat UNSIQ*, 9(1), 24-37. <https://doi.org/10.32699/ppkm.v9i1.2208>
- Martanto, M., & Hayati, U. (2024). Pengelompokan Transaksi Penjualan Aksesoris Hp Dan Pulsa Dengan Metode K-Means Untuk Meningkatkan Strategi Pemasaran Di Toko Bagus Cellular. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 2838-2849. <https://doi.org/10.36040/jati.v8i3.9559>
- Mulyadi, T. A., & Purnomo, D. (2023). Optimasi Pelayanan Kapal Penumpang melalui Clustering Penumpang dengan Metode Silhouette Coefficient. *Edumatic: Jurnal Pendidikan Informatika*, 7(2), 217-226. <https://doi.org/10.29408/edumatic.v7i2.21067>
- Nugraha, H. S., Mutaqin, H., Fathah, A., & Juliane, C. (2023). Mengidentifikasi Strategi Promosi pada Jasa Penjualan Saldo Digital menggunakan Pendekatan Clustering. *Edumatic: Jurnal Pendidikan Informatika*, 7(1), 11-19. <https://doi.org/10.29408/edumatic.v7i1.7385>

- Nurzahputra, A., Pranata, A. R., & Puwinarko, A. (2017). Decision Support System for Football Players Lineup Selection using Fuzzy Multiple Attribute Decision Making and K-Means Clustering Methods. *Jurnal Teknologi Dan Sistem Komputer*, 5(3), 106–109. <https://doi.org/10.14710/jtsiskom.5.3.2017.106-109>.
- Qirom, D. S., Faqih, A., & Dwilestari, G. (2024). Implementasi Algoritma K-Means Untuk Klasterisasi Pasien Hipertensi Berdasarkan Karakteristik Pasien. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 2056-2063. <https://doi.org/10.36040/jati.v8i2.8314>
- Rahman, F. D., Mulki, M. I. Z., & Taryana, A. (2024). Clustering dan klasifikasi data cuaca Cilacap dengan menggunakan metode K-Means dan Random Forest. *Jurnal SINTA: Sistem Informasi dan Teknologi Komputasi*, 1(2), 90-97. <https://doi.org/10.61124/sinta.v1i2.15>
- Salam, A., Adiatma, D., & Zeniarja, J. (2020). Implementasi Algoritma K-Means Dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS. *JOINS* 5(1), 62–68. <https://doi.org/10.33633/joins.v5i1.3350>
- Shelly, Z., Burch, R. F., Tian, W., Strawderman, L., Piroli, A., & Bichey, C. (2020). Using K-means clustering to create training groups for elite American football student-athletes based on game demands. *International Journal of Kinesiology and Sports Science*, 8(2), 47-63. <https://doi.org/10.7575//aiac.ijkss.v.8n.2p.47>
- Sulastri, H., Mubarok, H., & Iasha, S. S. (2021). Implementasi Algoritma Machine Learning Untuk Penentuan Cluster Status Gizi Balita. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, 5(2), 184-191. <https://doi.org/10.30872/jurti.v5i2.6779>
- Wadanur, A., & Sari, A. A. (2022). Implementasi Algoritma Apriori dan FP-Growth pada Penjualan Spareparts. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 107-115. <https://doi.org/10.29408/edumatic.v6i1.5470>