

KLASIFIKASI KARAKTER PEMAIN TUNGGAL PUTRA BULU TANGKIS DUNIA BERDASARKAN DATA KOMPETISI BWF MENGGUNAKAN ALGORITMA K-MEANS

Triadi Kurniawan¹, Ratih Kusuma Millennium²

¹Universitas Raharja; Tangerang; (021) 5529692

²Universitas Negeri Yogyakarta; Yogyakarta; (0274) 586168

Keywords:

Bulu Tangkis, BWF, Clustering, Data Mining, K-Means;

Correspondent Email:

triadi@raharja.info

Abstrak. Penelitian ini bertujuan untuk melakukan klasifikasi karakteristik pemain tunggal putra bulu tangkis dunia menggunakan data kompetisi resmi dari Badminton World Federation (BWF) periode 2022-2025. Data sebanyak 190,886 baris diproses melalui tahapan penggabungan, pembersihan, konversi, serta transformasi data menjadi 3,407 profil pemain unik. Algoritma K-Means Clustering diterapkan untuk membagi pemain ke dalam klasifikasi berdasarkan atribut performa yakni peringkat dunia, poin turnamen, dan jumlah partisipasi. Berdasarkan metode Elbow, jumlah klasifikasi optimal yang ditemukan adalah empat ($k=4$). Hasil penelitian menunjukkan empat klasifikasi dengan tipe karakteristik pemain: Low Exposure Players (44.61%), Beginner Players (42.21%), Developing Players (11.56%), dan Elite Players (1.61%). Evaluasi kualitas klasifikasi menggunakan Silhouette Score menghasilkan nilai 0.519 yang menandakan pemisahan klasifikasi cukup baik, serta Davies-Bouldin Index sebesar 0.641 yang menunjukkan struktur data terpisah dengan jelas. Temuan ini memberikan gambaran peta kekuatan bulu tangkis global, di mana mayoritas atlet dunia masih terkendala minimnya jam terbang turnamen internasional.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. This study aims to classifying the characteristics of men's singles badminton players worldwide based on competition data obtained from the Badminton World Federation (BWF) tournaments (2022-2025). The dataset comprises 190,886 raw data points, which were processed step-by-step compiling, cleansing, converting, and transforming data into 3,407 unique player profiles. The K-Means Clustering algorithm was employed to group players based on performance attributes such as rank, points, and tournament participation. The Elbow Method determined that the optimal number of clusters is four ($k=4$). The results identified four clusters and the characteristics each clusters are : Low Exposure Players (44.61%), Beginner Players (42.21%), Developing Players (11.56%), and Elite Players (1.61%). Evaluation using the Silhouette Score yielded a value of 0.519, indicating good cluster separation, while the Davies-Bouldin Index (DBI) of 0.641 confirmed the clusters are well-structured. This classification provides a comprehensive mapping of global badminton talent, highlighting that the majority of players fall into the low exposure category due to limited tournament access.

1. PENDAHULUAN

Bulu tangkis merupakan salah satu cabang olahraga paling populer di dunia dengan struktur kompetisi yang sangat ketat di bawah naungan *Badminton World Federation* (BWF). Setiap tahunnya, ribuan atlet dari berbagai negara bertanding dalam turnamen untuk mengumpulkan poin demi memperbaiki peringkat dunia mereka. Seiring dengan berkembangnya digitalisasi olahraga, volume data statistik yang dihasilkan dari pertandingan-pertandingan ini menjadi sangat besar dan kompleks. Pemanfaatan *Big Data* dan teknik *Data Mining* kini menjadi krusial untuk menggali wawasan dari data yang ada guna mendukung pengambilan keputusan yang strategis, baik bagi pelatih maupun federasi olahraga [1].

Perkembangan teknologi dalam olahraga telah mendorong penerapan berbagai algoritma cerdas untuk analisis performa. Sejumlah studi telah berhasil memanfaatkan *Machine Learning* untuk kebutuhan prediksi yang spesifik, seperti memprediksi hasil pertandingan berdasarkan frekuensi aksi teknis *badminton* tertentu misalnya *smash*, *net shot*, *drop shot*, dan lain-lain [2], membuat model prediksi taktis untuk sektor tunggal putri, guna menentukan strategi optimal selama pertandingan [3], hingga memodelkan pelatihan teknik dasar menggunakan *computer vision* [4], [5]. Di sisi lain, untuk kebutuhan klasifikasi data tanpa label sebelumnya, algoritma *K-Means Clustering* menjadi salah satu metode yang paling andal dan efisien [6]. Fleksibilitas dan validitas metode ini dalam menangani data numerik multidimensi telah teruji dalam berbagai domain, mulai dari klasifikasi prioritas logistik [7] hingga pemetaan karakteristik wilayah [8]. Keberhasilan penerapan metode ini di bidang lain menunjukkan potensi besar untuk diadopsi dalam menganalisis tingkat kemampuan atlet.

Meskipun eksplorasi teknologi dalam bulu tangkis terus berkembang, literatur yang ada saat ini masih memiliki keterbatasan fokus, di mana mayoritas analisis performa cenderung terpusat pada pemain level *elite* atau peringkat atas dunia, sebagaimana dilakukan oleh Hoffmann dan Vogt (2024) yang meneliti karakteristik pertandingan secara mendalam namun terbatas pada level tertinggi [9]. Sementara itu, pemetaan karakteristik ribuan

pemain lainnya yang berada di level menengah (*developing*) dan pemula (*beginner*) masih sangat jarang dilakukan. Ketiadaan klasifikasi yang komprehensif ini menyulitkan pemahaman mengenai peta distribusi kekuatan atlet secara global, padahal mayoritas populasi pemain justru berada di luar kelompok *elite*.

Penelitian ini bertujuan untuk menjembatani keterbatasan tersebut dengan melakukan klasifikasi karakteristik pemain tunggal putra dunia secara menyeluruh menggunakan data kompetisi BWF periode 2022-2025. Penerapan algoritma *K-Means* pada dataset berskala global seperti pada penelitian ini dapat memberikan pendekatan objektif untuk memetakan pemain ke dalam klaster-klaster performa yang berbeda. Hasil klasifikasi ini diharapkan dapat memberikan wawasan baru mengenai struktur piramida prestasi bulu tangkis dunia, serta dapat menjadi rujukan strategis bagi pengembangan atlet yang lebih merata di masa depan terkhusus pada sektor tunggal putra.

2. TINJAUAN PUSTAKA

2.1. Analisis Data Olahraga (Sports Analytics)

Pemanfaatan *Machine Learning* dalam olahraga telah berkembang pesat. Di bulu tangkis, data digunakan tidak hanya untuk statistik skor, tetapi juga untuk menganalisis pergerakan pemain dan akurasi pukulan [2]. Studi terbaru tahun 2024 bahkan menggunakan jaringan syaraf tiruan untuk memprediksi hasil pertandingan berdasarkan metrik visual atlet [5]. Hal ini menunjukkan bahwa data kompetisi menyimpan informasi berharga yang dapat digali lebih dalam.

2.2. Algoritma K-Means Clustering

K-Means adalah algoritma klasifikasi *non-hierarki* yang membagi data ke dalam sejumlah k klaster atau kelompok. Algoritma ini sangat populer karena sederhana dan cepat dalam mengolah data berjumlah besar [10]. Prinsip kerjanya adalah mencari pusat kelompok (*centroid*) dan menarik data-data di sekitarnya yang memiliki jarak terdekat ke pusat tersebut.

Penerapan *K-Means* telah terbukti efektif di berbagai bidang. Dalam jurnal JITET (2024), Fitriyani et al. menggunakan metode ini untuk mengelompokkan persediaan obat, membuktikan bahwa *K-Means* mampu

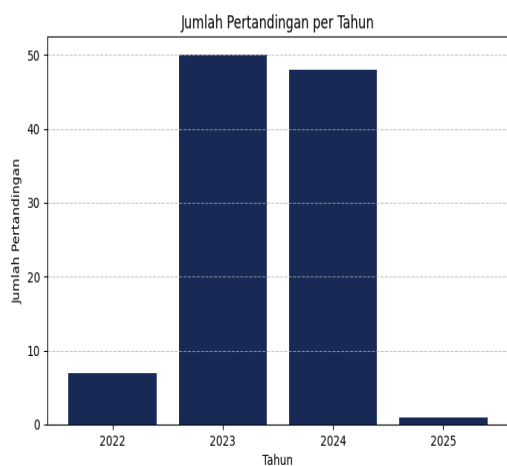
memisahkan item prioritas tinggi dan rendah dengan baik menggunakan evaluasi *Silhouette Score* [7]. Metode serupa juga diterapkan oleh Salsabila et al. (2024) untuk pemetaan wilayah, yang menunjukkan fleksibilitas algoritma ini untuk berbagai jenis data numerik [8]. Relevansi metode ini dengan penelitian olahraga juga dikonfirmasi oleh studi Mastroperris (2025) yang menggunakan *K-Means* untuk mengelompokkan performa atlet baseball di MLB [11].

3. METODE PENELITIAN

3.1. Pengumpulan Data (Data Collecting)

Data penelitian diambil dari repositori kompetisi *BWF World Tour* mulai 10 November 2022 hingga 6 Januari 2025 dengan komposisi pertandingan per tahunnya adalah sebagai berikut :

- Tahun 2022: 7 pertandingan
- Tahun 2023: 50 pertandingan
- Tahun 2024: 48 pertandingan
- Tahun 2025: 1 pertandingan.



Gambar 1. Distribusi Data Pertandingan 2022-2025

Berdasarkan gambar 1 ditunjukkan bahwa penelitian ini memiliki 106 *dataframe* yang merepresentasikan jumlah pertandingan selama kurun waktu 2022 sampai dengan 2025. *Dataframe* tersebut merupakan informasi peringkat pemain bulutangkis yang bersifat dinamis dan bergantung waktu, di mana setiap baris merepresentasikan kondisi peringkat

pemain pada periode tertentu. Perubahan nilai dalam *dataframe* terjadi akibat pembaruan hasil pertandingan, masa berlaku poin, serta dinamika performa pemain lain. Sehingga, *dataframe* ini mencerminkan perubahan longitudinal performa pemain dari waktu ke waktu. Jika seluruh data peringkat pemain dijumlah, maka total baris data sebanyak 190,886 baris.

Dataset disusun dalam struktur folder berdasarkan tanggal kompetisi, dengan format file yang konsisten pada setiap tanggal. Setiap file memuat data peringkat dan performa pemain pada waktu tertentu.

Adapun atribut data / variabel dalam penelitian ini adalah sebagai berikut :

Tabel 1. Atribut *Data Clustering*

Atribut	Deskripsi
player_id	Identitas unik pemain
name	Nama pemain
country	Negara asal pemain
continent	Benua asal pemain
rank	Peringkat dunia pemain
previous_rank	Peringkat dunia sebelumnya
points	Poin BWF pemain
number of tournament	Jumlah turnamen yang diikuti
competition_date	Tanggal kompetisi

3.2. Praproses Data (Data Preprocessing)

Praproses data melalui tahapan penggabungan, pembersihan, konversi, serta transformasi data guna mendapatkan data yang komprehensif dalam proses analisis. Berikut langkah-langkah praproses yang dilakukan :

3.2.1. Penggabungan Data

Tahap pertama dilakukan penggabungan data dengan mengumpulkan seluruh file kompetisi menjadi satu *dataframe* sehingga menghasilkan 190,886 baris. Selanjutnya, untuk memperoleh data tiap pemain maka dilakukan *grouping* data berdasarkan atribut *player_id*, *name*, *country*, dan *continent*, sehingga diperoleh data sebanyak 3,547 pemain.

3.2.2. Penanganan Missing Value

Data dengan nilai kosong (*NaN*) pada atribut numerik yang menjadi penentu model yaitu *rank*, *previous_rank*, *points*, dan *number_of_tournaments* penting untuk dihapus. Pada kasus ini setelah pembersihan diperoleh data sebanyak 3,528 pemain.

3.2.3. Penanganan Duplikasi Data

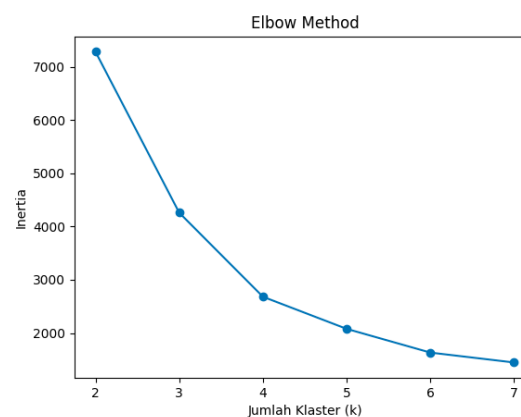
Setelah dilakukan penggabungan *dataframe*, masih ditemukan data yang belum unik (satu pemain dapat muncul pada banyak tanggal kompetisi), sehingga perlu dilakukan pembersihan data duplikat dengan melakukan *grouping data* berdasarkan *player_id* dan mengambil nilai rata-rata (*mean*) untuk setiap pemain pada atribut-atribut numerik seperti *rank*, *previous_rank*, *points*, dan *number_of_tournaments*. Setelah dilakukan penanganan duplikasi data, maka dataset akhir berjumlah 3,407 pemain.

3.2.4. Konversi dan Transformasi Data

Tahapan ini bertujuan menyesuaikan format dan skala data agar siap dimodelkan secara matematis. Pertama, dilakukan konversi tipe data di mana kolom *competition_date* diubah ke format *datetime*, sedangkan atribut numerik utama (*rank*, *previous_rank*, *points*, dan *number_of_tournaments*) dipastikan bertipe *float* atau *integer*. Selanjutnya, dilakukan transformasi data menggunakan teknik *StandardScaler* dari pustaka *Scikit-Learn*. Mengingat adanya perbedaan skala yang mencolok antar atribut, misalnya *points* memiliki rentang puluhan ribu sementara *rank* hanya satuan. Normalisasi ini penting untuk menstandarisasi fitur agar memiliki rata-rata 0 dan standar deviasi 1. Proses ini mencegah atribut dengan nilai besar mendominasi perhitungan jarak Euclidean pada algoritma *K-Means* [12].

3.3. Penentuan Jumlah Kluster (*k*)

Penentuan jumlah kluster atau kelompok yang paling ideal adalah dengan menggunakan *Metode Elbow*. Metode ini melihat penurunan nilai inersia (*error*). Proses ini dilakukan dengan menghitung nilai *Inertia* atau *Sum of Squared Errors (SSE)* yaitu jumlah kuadrat jarak antara setiap titik data dengan pusat klasternya (*centroid*). Berdasarkan implementasi kode, perhitungan inersia dilakukan secara iteratif dengan menguji jumlah kluster (*k*) mulai dari 2 hingga 7. Berikut adalah grafik hasil penentuan jumlah kluster pada penelitian ini :

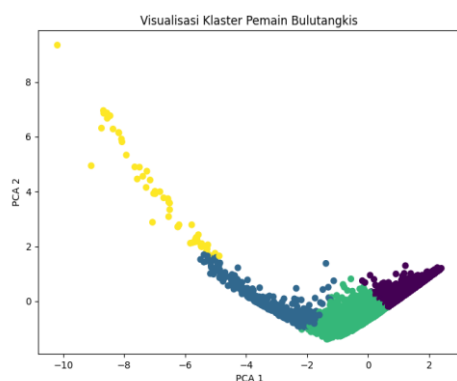


Gambar 2. Elbow Method

Berdasarkan gambar 2 hasil penentuan jumlah kluster tersebut, titik "siku" atau patahan yang paling tajam terjadi pada $k=4$, sehingga disimpulkan untuk membagi pemain ke dalam 4 kelompok atau kluster.

3.4. Teknik Visualisasi Data

Untuk mempermudah interpretasi hasil *clustering*, penelitian ini menggunakan teknik reduksi dimensi *Principal Component Analysis (PCA)*. Data multidimensi disederhanakan menjadi dua komponen utama (*principal components*) agar dapat divisualisasikan dalam bentuk grafik *scatter plot* 2 dimensi, sehingga pola penyebaran dan pemisahan antar kluster dapat diamati secara visual. Berikut adalah hasil visualisasi *clustering* yang ada :



Gambar 3. Visualisasi Kluster dengan PCA

Berdasarkan plot PCA pada Gambar 3 terlihat 4 kluster yang terpisah secara natural. Ini menandakan *K-Means Valid*. Berdasarkan plot tersebut, sumbu-x atau PCA 1 (garis horizontal) merepresentasikan kualitas pemain (kombinasi rank kecil + poin tinggi dari kiri ke kanan) serta sumbu-y PCA 2 (garis vertikal) merepresentasikan eksposur dan konsistensi turnamen.

3.5. Evaluasi Model

Kualitas hasil klasifikasi diuji menggunakan dua parameter:

- 1) **Silhouette Score:** Mengukur seberapa dekat seorang pemain dengan kelompoknya sendiri dibandingkan dengan kelompok lain. Nilai mendekati 1 berarti klasifikasi sangat baik [13]
- 2) **Davies-Bouldin Index (DBI):** Mengukur seberapa terpisah antar kelompok. Semakin kecil nilainya, semakin baik pemisahan antar kelompoknya [14]

4. HASIL DAN PEMBAHASAN

4.1. Hasil Clustering

Penerapan algoritma *K-Means* pada 3,407 pemain menghasilkan pembagian 4 kluster yang mana tiap kluster merepresentasikan kategori karakter pemain tunggal putra bulutangkis dunia. Kluster-kluster tersebut adalah sebagai berikut:

Tabel 2. Statistik Kluster

Kluster	Kategori	Jumlah Pemain	Persentase	Rata-Rata Rank	Rata-Rata Poin
0	Low Exposure	1,520	44.61%	1,461.13	122.35
1	Developing	394	11.56%	294.87	8,240.41
2	Beginner	1,438	42.21%	812.84	989.12
3	Elite	55	1.61%	31.20	47,692.59

4.2. Analisis Karakteristik Kluster

- 1) **Kluster Elite Player (1.61%):** Ini adalah kelompok eksklusif berisi 55 pemain top dunia seperti Viktor Axelsen dan Jonatan Christie. Mereka memiliki rata-rata poin sangat tinggi (>47,000) dan peringkat dunia 1-30 besar. Konsistensi mereka sangat tinggi karena rutin mengikuti turnamen besar.
- 2) **Kluster Developing Player (11.56%):** Berisi pemain seperti Kalle Koljonen (Finlandia) atau Mithun Manjunath (India). Mereka adalah pemain kompetitif yang sedang meniti karir menuju level elite, dengan rata-rata peringkat 294. Mereka cukup aktif bertanding namun belum mencapai poin maksimal.
- 3) **Kluster Beginner Player (42.21%):** Berisi pemain muda atau pemain dari negara berkembang bulu tangkis. Contohnya Abhinav Manota (New Zealand). Peringkat rata-rata mereka di kisaran 800-an dengan poin di bawah 1,000.
- 4) **Kluster Low Exposure Player (44.61%):** Ini adalah kelompok terbesar. Pemain di sini, seperti Carin Toudonou (Benin), memiliki peringkat di atas 1,000 dan poin sangat rendah (<150). Istilah "Low Exposure" digunakan karena rendahnya poin mereka kemungkinan besar disebabkan oleh minimnya kesempatan mengikuti turnamen internasional, bukan semata-mata karena kurang bakat teknis.

4.3. Evaluasi Model

Evaluasi menunjukkan nilai *Silhouette Score* sebesar 0.519. Nilai tersebut berada di atas 0.5 sehingga dapat dinilai sebagai struktur kluster yang reasonable atau cukup kuat [15]. Nilai *Davies-Bouldin Index* (DBI) sebesar 0.641 juga mengonfirmasi bahwa jarak antar kelompok cukup jauh yang berarti karakteristiknya cukup berbeda, terutama memisahkan kelompok *elite* dari kelompok lainnya secara signifikan [16]. Visualisasi data memperlihatkan adanya kesenjangan (*gap*) prestasi yang sangat lebar antara 55 pemain *elite* dengan ribuan pemain lainnya.

5. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan mengenai klasifikasi karakteristik pemain tunggal putra bulu tangkis dunia menggunakan algoritma *K-Means*, dapat ditarik beberapa kesimpulan sebagai berikut:

- 1) Penerapan algoritma *K-Means* pada 3,407 data pemain berhasil mengidentifikasi empat klasifikasi karakteristik yang berbeda, yaitu:
 - Kluster *Elite* (1.61%): Kelompok terkecil dengan performa puncak dan partisipasi turnamen tertinggi.
 - Kluster *Developing* (11.56%): Pemain kompetitif yang sedang berkembang menuju level elit.
 - Kluster *Beginner* (42.21%): Pemain pemula atau junior dengan poin relatif rendah.
 - Kluster *Low Exposure* (44.61%): Kelompok terbesar yang didominasi oleh pemain dengan peringkat rendah akibat minimnya jumlah turnamen yang diikuti.
- 2) Hasil pemetaan menunjukkan adanya kesenjangan yang signifikan dalam ekosistem bulu tangkis dunia. Dominasi kluster *Low Exposure* mengindikasikan bahwa hambatan utama bagi separuh populasi atlet dunia bukanlah sekadar kualitas teknis, melainkan kurangnya jam terbang atau kesempatan berkompetisi di level internasional.
- 3) Algoritma *K-Means* terbukti efektif dan valid untuk melakukan segmentasi atlet berdasarkan data kompetisi. Hal ini

dibuktikan dengan nilai evaluasi *Silhouette Score* sebesar 0.519 yang menunjukkan pemisahan kluster yang cukup baik, serta nilai *Davies-Bouldin Index* (DBI) sebesar 0.641 yang menandakan struktur data terkelompok dengan jelas.

- 4) Saran Pengembangan untuk penelitian selanjutnya, disarankan untuk:
 - Menambahkan variabel lain yang lebih spesifik seperti usia pemain, durasi pertandingan, atau statistik kemenangan *head-to-head* untuk mendapatkan profil karakteristik yang lebih mendalam.
 - Menerapkan metode *clustering* lain (seperti DBSCAN atau *Hierarchical Clustering*) sebagai pembanding untuk menguji konsistensi hasil klasifikasi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Raharja yang telah memfasilitasi penelitian ini, serta komunitas open source yang memungkinkan akses terhadap data BWF untuk tujuan akademik.

DAFTAR PUSTAKA

- [1] S. García et al., "Big data preprocessing: methods and prospects for sports analytics" *Big Data Analytics*, vol. 9, no. 1, 2023.
- [2] Z. Zhang, "Predicting badminton outcomes through machine learning and technical action frequencies" *Scientific Reports*, vol. 14, no. 1, p. 12345, 2025. [Online]. Available: <https://doi.org/10.1038/s41598-025-87610-7>
- [3] M. Suyal and S. Sharma, "Research on *K-Means* Clustering Optimization Algorithm Based on Machine Learning" *Journal of Artificial Intelligence and Systems*, vol. 6, no. 1, pp. 20-26, 2024. [Online]. Available: <https://doi.org/10.33969/AIS.2024060103>
- [4] H. Wang, "Machine Learning-Based Badminton Basic Technique Training Model" *International Journal of Human Movement and Sports Sciences*, vol. 12, no. 6, pp. 1023-1031, Nov. 2024. [Online]. Available: <https://doi.org/10.13189/saj.2024.120601>
- [5] B. T. Teo et al., "Predicting Shot Accuracy in Badminton Using Quiet Eye Metrics and Neural Networks," *Applied Sciences*, vol. 14, no. 21, p. 9906, 2024. [Online]. Available: <https://doi.org/10.3390/app14219906>

- [6] Y. Yuan, Y. Wang, K. Yang, and Y. Bin, "Prediction model and technical and tactical decision analysis of women's badminton singles based on machine learning," *PLoS ONE*, vol. 19, no. 11, p. e0312801, 2024. [Online]. Available: <https://doi.org/10.1371/journal.pone.0312801>
- [7] D. Fitriyani, M. Jajuli, and G. Garno, "Implementasi Algoritma K-Means Untuk Klasterisasi Dalam Pengelolaan Persediaan Obat (Studi Kasus: Apotek Naza)," *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, vol. 12, no. 3, pp. 2060-2068, 2024. [Online]. Available: <https://doi.org/10.23960/jitet.v12i3.4921>
- [8] F. Salsabila, T. Ridwan, and H. H., "Analisa Volume Penyebaran Sampah Di Karawang Menggunakan Algoritma K-Means Clustering" *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, vol. 12, no. 2, pp. 1309-1316, 2024. [Online]. Available: <https://doi.org/10.23960/jitet.v12i2.4226>
- [9] D. Hoffmann and T. Vogt, "Does a decade of the rally-point scoring system impact the characteristics of elite badminton matches?" *International Journal of Performance Analysis in Sport*, vol. 24, no. 2, pp. 105-118, 2024. [Online]. Available: <https://doi.org/10.1080/24748668.2023.2272121>
- [10] Y. Ramadhana and M. I. Jambak, "The Influence of Optimization of the k-Means Algorithm with Genetic Algorithm on the Results of High Dimension Data Clustering" *Indonesian Journal of Computer Science*, vol. 13, no. 1, p. 302, 2024.
- [11] M. Mastroperris, "Data visualization and K-means clustering in MLB: uncovering performance trends, payroll impact, and strategic insights" *Managing Sport and Leisure*, pp. 1-18, 2025. [Online]. Available: <https://doi.org/10.1080/23750472.2025.2557903>
- [12] R. Ishak, A. Bengnga, and R. T. Lopo, "Optimasi K-Means pada Clustering Penyakit Ibu Hamil Menggunakan Random Forest," *Jambura Journal of Electrical and Electronics Engineering*, vol. 6, no. 1, pp. 20-26, Jan. 2025. [Online]. Available: <https://ejournal.ung.ac.id/index.php/jjee/article/view/28374>
- [13] I. A. Rosyada and D. T. Utari, "Penerapan principal component analysis untuk reduksi variabel pada algoritma K-Means clustering" *Jambura Journal of Probability and Statistics*, vol. 5, no. 1, pp. 6-13, 2024. [Online]. Available: <https://doi.org/10.37905/jjps.v5i1.18733>
- [14] A. Mahiruna et al., "Enhancing Clustering Accuracy Using K-Means with Seeds Optimization," *Journal of Applied Informatics and Computing*, vol. 9, no. 5, pp. 10458-10465, 2025. [Online].
- [15] A. Fauzan, N. Suarna, I. Ali, and H. Susana, "Penerapan Algoritma K-Means Clustering untuk Meningkatkan Model Pengelompokan dan Kinerja Jaringan Wi-Fi Secara Optimal," *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, vol. 13, no. 2, pp. 594-603, 2025. [Online]. Available: <https://doi.org/10.23960/jitet.v13i2.6272>
- [16] M. F. A. Razak et al., "Cluster Analysis of Student Performance Data Using K-Means Algorithm with Davies-Bouldin Index Optimization," *Journal of Applied Technology and Information Systems*, vol. 5, no. 4, pp. 20-28, 2025.