International Journal of Human Movement and Sports Sciences 14(1): 1-11, 2026 http://www.hrpub.org
DOI: 10.13189/saj.2026.140101

Exploring Performance Patterns in Student Athletes: A Hierarchical Clustering Approach with Average Linkage

Syahril Bais[1], Padli[1,*], Denisha Aprilia Charles[2], Suci Nanda Sari[1], Heru Syarli Lesmana[1], Jeki Haryanto[1], Firunika Intan Cahyani[1]

[1]Faculty of Sports Sciences, Universitas Negeri Padang, Indonesia
[2]Faculty of Mathematics and Natural Sciences, Universitas Negeri Padang, Indonesia

Received August 24, 2025; Revised November 28, 2025; Accepted December 23, 2025

Cite This Paper in the Following Citation Styles

(a): *[1] Syahril Bais, Padli, Denisha Aprilia Charles, Suci Nanda Sari, Heru Syarli Lesmana, Jeki Haryanto, Firunika Intan Cahyani , "Exploring Performance Patterns in Student Athletes: A Hierarchical Clustering Approach with Average Linkage," International Journal of Human Movement and Sports Sciences, Vol. 14, No. 1, pp. 1 - 11, 2026. DOI: 10.13189/saj.2026.140101.*

(b): *Syahril Bais, Padli, Denisha Aprilia Charles, Suci Nanda Sari, Heru Syarli Lesmana, Jeki Haryanto, Firunika Intan Cahyani (2026). Exploring Performance Patterns in Student Athletes: A Hierarchical Clustering Approach with Average Linkage. International Journal of Human Movement and Sports Sciences, 14(1), 1 - 11. DOI: 10.13189/saj.2026.140101.*

Abstract This study examines the use of hierarchical clustering to identify key factors influencing athletic performance and to enhance coaching strategies at the West Sumatra Student Training Education Center (PPLP). A quantitative approach with total sampling was applied, involving 111 athletes across various sports. Data were collected using a validated questionnaire covering six variables: physical condition, social environment, athlete's character, training environment, coach's influence, and motivation. Reliability and validity were confirmed through exploratory factor analysis and Cronbach's Alpha ($\alpha > 0.70$). Data analysis included descriptive statistics, assumption testing, and hierarchical clustering with the average linkage method based on Euclidean distance. Results revealed five clusters: Cluster 1 (102 athletes) and Cluster 2 (6 athletes) were dominated by motivation; Cluster 3 (1 athlete) highlighted the coach's influence; Cluster 4 (1 athlete) emphasized the training environment; and Cluster 5 (1 athlete) underscored personal characters such as discipline and responsibility. Physical condition was significant mainly in Clusters 2 and 3. Overall, findings suggest that motivation, continuous coach support, and a conducive training environment are critical for optimizing performance. While physical condition remains important, it tends to interact with other factors rather than serve as the primary determinant.

Keywords Coaching Effectiveness, Student-Athlete Performance, Clustering Analysis, Performance Patterns

1. Introduction

The Training Center for Student Athletes serves as a platform for the sustainable development of excellence in sports across all regions of Indonesia. The talent development program, referred to as PPLP, is an initiative aimed at fostering the growth of sports in Indonesia, managed by the Regional Technical Implementation Unit (UPTD) of the Youth and Sports Office. The PPLP program functions as a central hub for the coaching and development of student athletes, with the objective of enhancing sports performance at the provincial level. This goal is achieved through a structured training regimen that encompasses systematic planning, a diverse array of training materials, and the implementation of appropriate training methodologies to optimize athlete performance.

The variety of sports offered at the student athlete development center possesses unique characteristics, with each sport presenting its own advantages and challenges. These characteristics necessitate a clustering approach to categorizing sports based on specific attributes, thereby facilitating easier analysis, development, and application across various domains, including athlete coaching, training methodologies, strategic development, and scientific research.

Clustering aims to optimize an athlete's performance achievements [1]. According to Shelly [2], clustering serves as an effective strategy for forming training groups for student-athletes within organizations focused on athletic development, particularly in environments with a large number of participants. This approach aims to enhance the foundational training of junior athletes. While clustering can elevate athletes' performance levels, its ongoing application may also lead to drawbacks, potentially restricting the development of diverse skills and thereby impeding their long-term growth [3]. According to Yang [4] and Fountain [5], the grouping method in sports is designed to categorize athletes based on their characteristics. When this method is applied in the context of sports psychology, elite athletes can be classified based on their behaviors, which are influenced by several key factors, including social, cultural, and economic elements [6].

The term "student athlete" commonly describes an individual who balances two primary responsibilities: those of a student and an athlete. It is essential for these dual roles to be managed in a harmonious manner to ensure optimal performance in both areas. This aligns with the perspectives of Dewi [7], and SANBERK [8], adolescent athletes face a unique challenge: they must juggle their responsibilities as students while fully committing to their sport. It is found that perfect balance is essential for their growth and success in both areas. Every young athlete thrives on support, and the environment around them plays a crucial role in their development [9]. From family encouragement to community involvement, and access to top-notch facilities and infrastructure, these elements are essential in ensuring the coaching process unfolds as it should. When all these pieces come together, youth athletes have the foundation they need to reach their full potential. The grouping method helps coaches identify issues related to athlete quality by examining intrinsic and extrinsic factors, facilitating the development of an effective athlete support strategy [10]. Clustering is not just about enhancing physical and psychological aspects; it is also an effective method for grouping athletes with similar characteristics [11]. This approach can improve the accuracy of screening protocols used to identify individuals at high or low risk of injury.

Cluster analysis is a method used to group objects or problems into relatively similar categories, known as clusters [12]. Items within each cluster exhibit high intra-cluster similarity while demonstrating significant dissimilarity from items in other clusters [13]. The average linkage method is a hierarchical clustering technique that calculates the inter-cluster distance based on the average distance between all pairs of objects from one cluster to all objects within another cluster [14]. This approach provides a more nuanced representation of cluster similarity by taking into account the collective distribution of points, rather than relying solely on the nearest or furthest individual points. According to Miller [15], there is speculation regarding the reasons for the formation of this grouping, particularly concerning the involvement of coaches and academic support personnel in influencing the coaching strategies employed to facilitate student athletes in achieving their goals.

Speculation exists regarding why the grouping occurred, particularly due to the role of coaches and academic support personnel in guiding the coaching strategy of student-athletes in achieving success [16]. Clustering plays a vital role in optimizing athletes' talents [17]. Advancing knowledge and understanding of performance management is essential for elite sports organizations [18], [19], and clustering can be an effective solution for enhancing athlete performance and achievements. Research by McGarry et al. [20], and Grasdalsmoen et al. [21], highlights two approaches to cluster analysis in sports: exploratory and confirmatory. The exploratory approach utilizes hierarchical cluster analysis to gain a comprehensive understanding of movement patterns as a whole. The process begins with an initial object, followed by the individual grouping of subsequent objects based on a defined metric until all objects are integrated. This approach is valuable for identifying patterns within a collection of predicted movements or for discovering new categories that require further analysis.

Assessing physical abilities in sports is essential for determining and enhancing athletic performance across all disciplines [22]. Concerns regarding academic grouping in sports, particularly its effect on minority players, warrant academic and media scrutiny [23]. In achievement sports, specific physical attributes and postures are essential based on the requirements of the sport in question, such as sepaktakraw. In the game of sepaktakraw, an important anthropometric factor to consider is the ratio of height to leg length and body weight [24], [25]. Identifying latent profiles in athletes based on their mental and physical health is crucial for clustering elite athletes [26]. This helps coaches train strategies, techniques, and tactics to enhance athletes' performance [27].

In competitive sports, the primary objective of training is to optimally enhance performance, making athletes' physical abilities the fundamental factor in improving their overall capabilities [28]. In addition, according to the Self-Determination Theory [29], motivation is often considered the key to an athlete's success, which is related to the role of the coach as a model for the coach-athlete

relationship that influences the achievement of the athlete's psychological and psychological components [30]. Other studies also mention the relationship between physical condition, which interacts with motivation, training environment, and coach support as determining factors in improving athlete performance [31], [32]. Therefore, in the context of competitive sports, clustering has been used to identify training typologies, performance profiles, and motivational structures [33].

Clustering can serve as an effective coaching strategy to enhance athletes' performance. The clustering method allows for the identification of strengths and weaknesses within the sports coaching framework by categorizing athletes according to factors that influence their performance. This grouping is designed to help coaches and managers tailor training programs and coaching strategies, ultimately boosting motivation and optimizing athletes' achievements. Consequently, this study employed cluster analysis utilizing the average linkage method to identify groups of athletes based on the factors that contribute to their success.

2. Materials and Method

2.1. Research Design

This study employed a quantitative, applied research design to examine the effectiveness of clustering as a coaching strategy to enhance athletic performance among student-athletes. The research was conducted at the West Sumatra Student Training Education Center (PPLP), focusing on the identification and classification of factors that influence athlete achievement. The approach integrates both descriptive and inferential statistical techniques, culminating in a hierarchical cluster analysis using the average linkage method to categorize athletes based on performance-influencing variables.

2.2. Participants

A total sampling technique was utilized, whereby the entire population of interest was selected as the sample. The study included 111 student-athletes from various sports disciplines registered at PPLP West Sumatra. This method ensured a comprehensive representation of the population, eliminating sampling bias and enhancing the validity of the clustering process.

2.3. Procedure and Research Instrument

The primary instrument used for data collection was a structured questionnaire, developed based on a theoretical framework that identifies key factors influencing athlete performance. The questionnaire consisted of closed-ended items measured on a Likert scale, encompassing six core variables:

1. Athlete's Physical Condition (X1)
2. Social Environment (X2)
3. Athlete’s Character (X3)
4. PPLP (Student Center Training) Environment (X4)
5. Coach’s Influence (X5)
6. Motivation (X6)

Prior to deployment, the instrument underwent content validation by experts in sports science and pilot testing to assess clarity and comprehensibility. Construct validity was examined through factor analysis, and reliability was tested using Cronbach's Alpha to ensure internal consistency, with all scales achieving an alpha value > 0.70.

2.4. Data Collection Procedures

The data collection process involved the distribution of the validated questionnaire to the 111 athletes during scheduled training sessions at the PPLP facility. Athletes were briefed about the purpose of the research and were assured of confidentiality and anonymity. Responses were collected over a two-week period to ensure maximum participation. The collected data were then coded and entered into IBM SPSS Version 24 for analysis.

2.5. Statistical Analysis

The data analysis was conducted in several stages:

1. Descriptive Statistics:
Frequencies and percentages were calculated to describe the demographic distribution of athletes by gender and type of sport.

2. Validity and Reliability Testing:
   a. Construct validity was assessed using exploratory factor analysis (EFA).
   b. Reliability of the instrument was measured using Cronbach’s Alpha.

3. Assumption Testing:
   a. Multicollinearity Test: Tolerance and Variance Inflation Factor (VIF) values were analyzed to detect multicollinearity among the independent variables. Acceptable thresholds were set at Tolerance > 0.10 and VIF < 10.
   b. Data Sufficiency Test: The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett’s Test of Sphericity were used to ensure the data were suitable for cluster analysis. KMO values greater than 0.5 and significant Bartlett’s test results ($p < 0.05$) indicated data adequacy.

4. Hierarchical Cluster Analysis:
Upon meeting the necessary assumptions, hierarchical cluster analysis was conducted using the average linkage (between groups) method. The steps included:
   a. **Euclidean Distance Calculation**: Measured the similarity or dissimilarity between athletes.

b. **Cluster Formation**: Athletes were grouped using the average linkage algorithm, and the clustering process was visualized through a dendrogram.
c. **Cluster Membership Determination**: The number of clusters and their respective members were identified.
d. **Cluster Interpretation**: Each cluster was analyzed to identify the dominant variables influencing athlete performance within the group.

All statistical analyses were performed using IBM SPSS Statistics Version 24, and findings were interpreted based on empirical patterns derived from the cluster groupings.

3. Results

3.1. Data Analysis by Gender

The data presented in Table 1 illustrates the gender distribution of the 111 athletes who participated in the study. Of the total respondents, 80 were male athletes, representing 72% of the sample, while 31 were female athletes, accounting for the remaining 28%. This indicates a significant predominance of male participants in the athlete population at the West Sumatra Student Training Education Center (PPLP). The gender disparity observed may reflect the overall enrollment trends or gender-based participation preferences in competitive sports within the institution. The data was collected through a questionnaire completed by 111 athletes, who can be generally grouped as follows:

Table 1. Data Analysis by Gender

No	Gender	Frequency	
		Absolute	*Relative*
1	Woman	31	28%
2	Man	80	72%
Sum		111	100%

3.2. Data Analysis by Sports

As shown in Table 2, the athletes in this study are involved in 12 different sports disciplines. The highest representation comes from football, with 18 athletes (16%), followed by sepaktakraw with 15 athletes (14%), and pencak silat with 13 athletes (12%). Other sports with notable participation include wrestling (11%), floor gymnastics (9%), and athletics (8%). Disciplines with lower representation include bike racing and karate (each with 5 athletes, 5%), as well as judo and archery (each with 4 athletes, 4%). Both taekwondo and boxing have an equal number of participants, with 8 athletes each (7%). This distribution highlights that while football and sepaktakraw dominate the athlete composition at PPLP West Sumatra, a wide variety of sports are supported and actively pursued, reflecting the institution's diverse athletic development program.

Table 2. Data Analysis by Sports

No	Sports	Frequency	
		Absolute	*Relative*
1	Athletics	9	8%
2	Bike Racing	5	5%
3	Wrestling	12	11%
4	Judo	4	4%
5	Karate	5	5%
6	Archery	4	4%
7	Pencak Silat	13	12%
8	Floor gymnastics	10	9%
9	Football	18	16%
10	Sepaktakraw	15	14%
11	Taekwondo	8	7%
12	Boxing	8	7%
Sum		111	100%

3.3. Athlete Clustering Based on Performance-Influencing Variables

As shown in Table 3, the clustering analysis grouped the 111 athletes into five distinct clusters based on the variables that most significantly affect their performance.

- Cluster 1 is the largest, comprising 102 athletes, where motivation (X6) is the most influential factor, followed by coach (X5), athlete character (X3), social environment (X2), PPLP environment (X4), and finally physical condition (X1).
- Cluster 2 consists of 6 athletes, with a similar pattern to Cluster 1, except that athlete's physical condition (X1) appears before PPLP environment (X4).
- Cluster 3, Cluster 4, and Cluster 5 each contains 1 athlete, reflecting unique profiles: (1) Cluster 3 prioritizes coach (X5), followed by physical condition, athlete character, and other variables. (2) Cluster 4 places PPLP environment (X4) as the top factor, indicating institutional influence as the primary performance driver for this individual. (3) Cluster 5 emphasizes athlete character (X3), with less importance on motivation and physical condition.

These results highlight that motivation is the most dominant variable for the majority of athletes, while coaching and athlete character also play significant roles. However, for some individuals, institutional or environmental factors emerge as primary influences, suggesting the need for personalized approaches in athlete development programs.

3.4. Assumption Test

3.4.1. Multicollinearity Test

A multicollinearity test is conducted to determine whether there is a strong correlation between independent variables. In this test, the researcher used IBM SPSS 24, with the criteria that the tolerance value should be greater than 0.10 and the Variance Inflation Factor (VIF) should be less than 10.00.

Table 4 shows that the tolerance value is greater than 0.10, and the VIF value is less than 10.00. This indicates that there is no multicollinearity present, meaning there is not a strong correlation between the independent variables.

3.4.2. Data Sufficiency Test

The Kaiser–Meyer–Olkin (KMO) and Bartlett's Test of Sphericity were applied to assess the adequacy and suitability of the data for cluster analysis [34]. The KMO value of 0.817 indicates meritorious sampling adequacy, meaning that correlations among variables are strong enough for multivariate analysis. Bartlett's Test yielded a significant result (p = 0.000), confirming that the correlation matrix is not an identity matrix and that relationships among variables are meaningful. Together, these results justify that the data structure is appropriate for cluster analysis and can produce valid and interpretable groupings. The KMO test was conducted using SPSS 24, among others, as shown in Table 5.

Table 5 indicates that the KMO value is 0.803. Since this value is greater than 0.5 and the significance level is less than 0.05, we can conclude that the data adequately represents the population. This suggests that the data is suitable for cluster analysis and can be used for further research.

Table 3. Analysis of Athlete Clustering Data

Cluster	Number of Athletes	Variables Affecting Performance (Largest to Smallest)
Cluster 1	102	Motivation (x6), Coach (x5), Athlete Character (X3), Social Environment (X2), PPLP Environment (X4), Physical Condition (X1)
Cluster 2	6	Motivation (x6), Athlete Character (X3), Social Environment (X2), Athlete's Physical Condition (X1), PPLP Environment (X4)
Cluster 3	1	Coach (x5), Athlete's Physical Condition (X1), Athlete's Character (X3), Motivation (X6), Social Environment (X2), PPLP Environment (X4)
Cluster 4	1	PPLP Environment (X4), Coach (X5), Social Environment (X2), Motivation (X6), Athlete Character (X3), Athlete Physical Condition (X1)
Cluster 5	1	Athlete Character (X3), PPLP Environment (X4), Coach (X5), Social Environment (X2), Motivation (X6), Athlete's Physical Condition (X1)

Table 4. Multicollinearity Test

Variabel	Tolerance Value	Information	VIF	Information
X1	0,827	No Multicollinearity	1,209	No Multicollinearity
X2	0,437	No Multicollinearity	2,289	No Multicollinearity
X3	0,571	No Multicollinearity	1,751	No Multicollinearity
X4	0,605	No Multicollinearity	1,652	No Multicollinearity
X5	0,49	No Multicollinearity	2,039	No Multicollinearity
X6	0,742	No Multicollinearity	1,349	No Multicollinearity

Table 5. KMO Data Adequacy Test *and Barlett's Test*

KMO and Bartlett's Test		
Kaiser - Mayer - Olkin Measure of Sampling Adequacy		.817
Bartlett's Test of Sphericity	Approx. Chi-Square	209.188
	df	15
	Sig.	.000

3.5. Analysis of the Hierarchy Clustering - Using the Average Linkage Method

Distance of Euclidean

Euclidean distance tables serve as valuable tools for assessing the proximity or similarity between points, facilitating the data grouping process, and simplifying the identification of cluster centers. They also help in determining the maximum distance between points. In this case, there are 111 cases that were ultimately grouped into five clusters. The clustering process began with cases that had relatively close average distances. For example, case 32 was merged with case 111 to form Cluster 1. Case 86 was merged with case 110 to form Cluster 2, case 85 was merged with case 107 to form Cluster 3, and case 17 was merged with case 56 to form Cluster 4. Next, case 71 was added to Cluster 3 (Cluster 3: 85, 107, 71). Case 38 was then merged with Cluster 4 (Cluster 4: 17, 56, 38). Cluster 2 (86, 110) was combined with Cluster 3 (85, 107, 71), resulting in Cluster 2 (71, 85, 86, 107, 110). Case 18 was then added to Cluster 2 (Cluster 2: 18, 71, 85, 86, 107, 110). Subsequently, Cluster 2 (18, 71, 85, 86, 107, 110) was merged with Cluster 4 (17, 56, 38), producing Cluster 2 (17, 18, 38, 56, 85, 86, 107, 110). This cluster was later combined with Cluster 1 (32, 111), forming Cluster 1 (17, 18, 32, 38, 56, 85, 86, 107, 110, 111) with an average Euclidean distance of approximately 10. At the same distance level (±10), other clusters were also identified. When further consolidated, within the average distance range of 15–20, five clusters were obtained, as shown in Tables 6 and 7.

3.6. Grouping Process with the Average Linkage Method

The average linkage method is an agglomerative or merging approach. After determining the Euclidean distance matrix, the next step involves merging into new groups or clusters. This approach represents a hierarchical clustering technique in which objects are progressively grouped according to their degree of similarity or distance. At each step, the cases with the smallest average distance are merged first, thereby forming a new group or cluster.

3.7. Determine the Number of Cluster Members

There is a clustering table that categorizes athletes and their respective sports into five distinct clusters. The analysis results, including the number of members in each cluster, can be found in Tables 6 to 8. This stage can be conducted using IBM SPSS 24.

Table 6. Clustering Process (Step by Step)

Step	Merging Process	Resulting Cluster
1	Case 32 + Case 111	Cluster 1 → (32, 111)
1	Case 86 + Case 110	Cluster 2 → (86, 110)
1	Case 85 + Case 107	Cluster 3 → (85, 107)
1	Case 17 + Case 56	Cluster 4 → (17, 56)
2	Case 71 added to Cluster 3	Cluster 3 → (85, 107, 71)
2	Case 38 added to Cluster 4	Cluster 4 → (17, 56, 38)
3	Cluster 2 (86, 110) + Cluster 3 (85, 107, 71)	Cluster 2 → (71, 85, 86, 107, 110)
3	Case 18 added to Cluster 2	Cluster 2 → (18, 71, 85, 86, 107, 110)
4	Cluster 2 (18, 71, 85, 86, 107, 110) + Cluster 4 (17, 56, 38)	Cluster 2 → (17, 18, 38, 56, 85, 86, 107, 110)
5	Cluster 2 (17, 18, 38, 56, 85, 86, 107, 110) + Cluster 1 (32, 111)	Cluster 1 → (17, 18, 32, 38, 56, 85, 86, 107, 110, 111)

*At this stage, the average Euclidean distance was approximately **10**, and other clusters also existed at this level.

Table 7. Final Clusters (Consolidated at Distance 15–20)

Cluster	Members
Cluster 1	32, 111, 17, 56, 38, 86, 110, 85, 107, 71, 18, 34, 53, 26, 30, 15, 95, 8, 19, 22, 5, 9, 37, 3, 21, 23, 89, 45, 6, 7, 66, 105, 77, 33, 92, 59, 64, 98, 25, 94, 36, 88, 108, 65, 104, 80, 76, 63, 103, 2, 1, 93, 79, 10, 67, 106, 96, 87, 54, 83, 99, 90, 97, 16, 70, 84, 57, 74, 40, 101, 29, 48, 68, 72, 61, 49, 81, 55, 12, 24, 42, 51, 11, 69, 52, 58, 28, 27, 75, 78, 62, 39, 100, 73, 60, 102, 91, 109, 31, 82, 14, 47
Cluster 2	13, 43, 44, 50, 46, 4
Cluster 3	20
Cluster 4	41
Cluster 5	35

Table 8. Cluster Member Based on Sports Disciplines

Cluster	Number of Members	Sports Disciplines Involved	Interpretation
Cluster 1	102	Athletics, Cycling, Wrestling, Judo, Karate, Archery, Pencak Silat, Floor Gymnastics, Football, Sepaktakraw, Taekwondo, Boxing	The majority group has diverse sports disciplines but relatively similar characteristics.
Cluster 2	6	Athletics, Cycling, Pencak Silat	A small subgroup, dominated by pencak silat, is internally consistent but distinct from the main cluster.
Cluster 3	1	Wrestling	Highly unique characteristics; stands alone in its own cluster.
Cluster 4	1	Karate	Significantly different from other clusters; forms a single-member cluster.
Cluster 5	1	Pencak Silat	Although pencak silat appears in other clusters, this individual shows a distinctly different profile.
Cluster 6	1	Boxing	The boxing athlete demonstrates uniquely different traits compared to all other clusters.

3.8. Cluster Interpretation

Once the number of clusters and their members has been established, the next step is to create profiles for each cluster. This profiling aims to identify the characteristics of each group. Table 9 displays the average calculations for each cluster.

According to Table 9, the characteristics of each cluster can be understood through the average values of each cluster, with the following interpretation:

1. Cluster 1 consists of 102 athletes, with the motivation variable (X6) demonstrating the strongest influence on achievement, scoring 3.463 out of 4.00. This is followed by the coach variable (X5), which received a score of 3.2 out of 4.00. The athlete character variable (X3) scored 3.163 out of 4.00, while the social environment variable (X2) achieved a score of 3.061 out of 4.00. The PPLP environmental variable (X4) received a score of 2.897 out of 4.00, and finally, the physical condition of the athlete (X1) was assessed at 2.765 out of 4.00 in terms of its impact on achievement.
2. Cluster 2 has 6 athletes, where the motivation variable (X6) has the greatest influence on achievement, which is 3.667 out of 4.00 on achievement, then followed by the coach variable (X5) of 3.650 out of 4.00, the athlete character variable (X3) of 3.450 out of 4.00, the social environment variable (X2) of 3.375 out of 4.00, the physical condition variable of athletes (X1) of 3.367 out of 4.00, and finally, the PPLP environmental variable (X4) had an effect of 3.175 out of 4.00 on achievement.
3. Cluster 3 has 1 athlete member, where the coach variable (X5) gives the greatest influence of 3.250 out of 4.00 on achievement, followed by the athlete's physical condition variable (X1) of 3.167 out of 4.00, and the athlete character variable (X3) and the motivation variable (X6) both exert an influence of 3.00 out of 4.00 on achievement. The social environment variable (X2) influenced 2.875 out of 4.00, and finally, the PPLP environmental variable (X4) affected 2.75 out of 4.00 on achievement.
4. Cluster 4 comprises one athlete member, where the PPLP environmental variable (X4) exerts the greatest influence on achievement, scoring 3.875 out of 4.00. This is followed by the coach variable (X5), which has a score of 3.750 out of 4.00. The social environment variable (X2) demonstrates an influence of 3.375 out of 4.00, while the motivation variable (X6) scores 3.333 out of 4.00. The athlete character variable (X3) receives a score of 3.12 out of 4.00, and lastly, the athlete's physical condition variable (X1) has the least influence, with a score of 2.667 out of 4.00 on athlete.
5. In Cluster 5, there is one athlete member. The athlete character variable (X3) has the most significant influence on achievement, with a score of 3.875 out of 4.00. This is followed by the PPLP environmental variable (X4) at 3.5 out of 4.00, and the coach variable (X5) at 3.375 out of 4.00. Both the social environment variable (X2) and the motivation variable (X6) contribute equally, each with a score of 3.00 out of 4.00. Lastly, the athlete's physical condition variable (X1) has the least influence, with a score of 2.333 out of 4.00 regarding achievement.

Table 9. Cluster Insights

Variabel	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
X1	2,765	3,367	3,167	2,667	2,333
X2	3,061	3,375	2,875	3,375	3,000
X3	3,163	3,450	3,000	3,125	3,875
X4	2,897	3,175	2,750	3,875	3,500
X5	3,200	3,650	3,250	3,750	3,375
X6	3,463	3,667	3,000	3,333	3,000

4. Discussion

In this study, the research variables are categorized into two main groups: exogenous and endogenous variables. The exogenous variables include athletes' physical qualities, social environment, character, the PPLP environment, coaching perceptions, and motivation. Conversely, the endogenous variables pertain to athletes' achievements. The analysis indicates that the majority of athletes fall into cluster 1, comprising a total of 102 individuals. Among the factors influencing achievement, motivation emerges as the most significant, followed by support from coaches and the athletes' character. Notably, the physical condition of the athlete appears to exert the least influence on achievement outcomes. These findings underscore the critical role of internal motivation as a key driver for athletes striving to attain their goals within this group [35]. In addition, motivation accompanied by high self-discipline can predict positive behavior to support the life of athletes in the future [36]–[38].

The results of this study show that motivation is the most influential factor on athletes' achievements. Studies show that intrinsic motivation and self-efficacy strongly predict athletes' success compared to physical or environmental factors [39]. Meta-analytic findings also confirm that motivation-enhancing interventions significantly improve engagement and performance [40]. The previous study identified motivation as the dominant psychological predictor of peak performance in sport [41]. According to research by Freire et al. [33], cluster analysis of young athletes in Brazil also found that groups of athletes with high motivation and positive coach-athlete relationships showed better performance and team cohesion than clusters with low motivation.

In the second group, which consists of six athletes, motivation emerges as the primary factor driving achievements, followed closely by the support of coaches and the character of the athletes themselves. Unlike in the first cluster, physical condition plays a significant role and is more influential than the PPLP environment. Athletes in this group generally exhibit better physical fitness and are notably impacted by their personal motivation. The findings indicate that self-driven motivation has a substantial effect on athletes' physical condition. According to Trotter et al. [42], psychological abilities and motivation are key contributors to enhancing performance and well-being during both competitions and training.

In the third cluster, the athletes are notably reliant on coach support as the most influential factor in their success. Physical condition ranks as the second most important factor, while the PPLP environment has the least impact. This indicates that these athletes heavily depend on their coaches' direction and guidance to achieve their goals. A supportive environment positively contributes to the sustainability of their accomplishments. Coaches play a crucial role as facilitators, effectively guiding athletes toward their established targets [43]–[46]. The mental well-being of athletes reflects the relationship between them and their coaches, where a coach's ability to communicate with and understand athletes is a crucial element in the success of competitive sports [47]–[50].

In contrast to the other clusters, the fourth cluster identifies the PPLP environment as the most influential factor, followed closely by the support provided by trainers. The factor with the least impact is physical condition. Athletes within this cluster appear to be significantly affected by their training environment and the available facilities. The training environment plays a crucial role in determining athletes' performance. According to Larsen at al. [51], a healthy coaching environment encompasses coaches actively supporting the development and implementation of strategies that foster intrinsic motivation, enjoyment, and long-term participation among athletes. Social support is essential in sports coaching, as it significantly contributes to enhancing athletes' psychological capabilities, particularly in terms of emotional regulation and overall mental well-being [42]. Research from various studies indicates that enhancing the quality of athletes in talent development environments significantly impacts their progress. Therefore, it is essential for sports developers to prioritize the creation of a safe environment that fosters the psychological development of athletes [52].

In the fifth cluster, the athlete's character is the most significant factor influencing achievements. A notable secondary factor is the PPLP environment and the support from coaches, while the impact of physical condition is

comparatively minimal. This athlete possesses personal traits such as discipline and responsibility, which are essential for achieving success. The integration of sport into positive youth development has been a subject of research study [53]. This is because the benefits of sports go beyond physical strength; sports coaching has become an effective means of enhancing the life skills of the younger generation [54]–[56]. Theoretically, the influences of exogenous variables vary for each athlete, resulting in different responses and, consequently, varied achievement outcomes. The clustering derived from this study can inform the training process, enabling PPLP coaches and administrators to tailor their interventions based on these clusters, thereby maximizing athletes' performance potential.

5. Conclusions

This study confirms that motivation, coach support, and training environment are the main determinants of athlete success, while physical condition functions more as a supporting factor that interacts with psychological and social aspects. These findings imply that coaching strategies should not only focus on physical strengthening, but also integrate psychological approaches to foster intrinsic motivation, build constructive coach-athlete relationships, and create a conducive training ecosystem. For sports managers, these results highlight the importance of planning development programs that balance technical, mental, and environmental aspects so that athletes' potential can be optimally developed. In the future, similar studies can be expanded to various sports or different populations of athletes to test the consistency of the patterns of factors found, as well as through longitudinal designs to understand the dynamics of changes in these factors throughout an athlete's career. Thus, this study not only contributes empirically to the sports coaching literature but also offers a practical framework for coaches and administrators in formulating more adaptive, comprehensive, and long-term athlete development-oriented coaching strategies.

Although this study successfully identified key factors that influence athlete performance through hierarchical clustering, there are several limitations that should be noted. The use of total sampling in one PPLP has the potential to cause sampling bias, while determining the optimal number of clusters in the analysis is subjective and risks causing overfitting when clusters with very small members appear. Therefore, further research should involve a more diverse sample, apply alternative clustering methods, and consider longitudinal designs or qualitative approaches to gain a more comprehensive understanding of the dynamics of athlete performance.

Acknowledgements

The author would like to express sincere gratitude to **PPLP (Student Training Education Center)** for the invaluable support and facilitation provided throughout the course of this study. The contributions of PPLP, particularly in providing access to athletes, training facilities, and data resources, have been instrumental in the successful completion of this research. Special thanks are extended to the coaches, staff, and athletes in the PPLP program, whose participation, cooperation, and dedication greatly enriched the findings and allowed for a deeper understanding of the factors influencing athletic performance. The continued commitment of PPLP to fostering young athletic talent is both commendable and inspiring.

REFERENCES

[1] V. A. Calhoun, “Division I Student Athletes and the Experience of Academic Clustering,” *Thesis Calif. State Univ. Long*, pp. 1–160, 2012.

[2] Z. Shelly, R. F. V. Burch, W. Tian, L. Strawderman, A. Piroli, and C. Bichey, “Using K-means clustering to create training groups for elite american football student-athletes based on game demands,” *Int. J. Kinesiol. Sport. Sci.*, vol. 8, no. 2, pp. 47–63, 2020, doi: 10.7575//aiac.ijkss.v.8n.2p.47.

[3] M. Barth and A. Güllich, “Non-linear association of efficiency of practice of adult elite athletes with their youth multi-sport practice,” *J. Sports Sci.*, vol. 39, no. 8, pp. 915–925, 2021, doi: 10.1080/02640414.2020.1851900.

[4] D. Yang, J. Wang, J. He, and C. Zhao, “A clustering mining method for sports behavior characteristics of athletes based on the ant colony optimization,” *Heliyon*, vol. 10, no. 12, pp. 1–12, 2024, doi: 10.1016/j.heliyon.2024.e33297.

[5] J. J. Fountain and P. S. Finley, “Academic Majors of Upperclassmen Football Players in the Atlantic Coast Conference: An Analysis of Academic Clustering Comparing White and Minority Players,” *J. Issues Intercoll. Athl.*, vol. 2, pp. 1–13, 2009, doi: 10.4324/9781315721606-106.

[6] K. Rahayuni, “Where Indonesian Elite Athletes Come From ? The Role of Family and Physical Where Indonesian Elite Athletes Come From ?,” *Proceeding Int. Webinar Phys. Lit.*, pp. 14–21, 2021.

[7] N. S. Dewi, P. Jittanoon, and W. Wiroonpanich, “Understanding healthy body and dieting of youth athletes among javanese muslim parents,” *Food Res.*, vol. 4, pp. 1–5, 2020, doi: 10.26656/fr.2017.4(S3).S01.

[8] İ. SANBERK, C. TÜRKERİ, and B. PARSAK, “Investigation of Personal and Social Sports Identities of Elite National Karate Players via Self Identity Chart,” *Spor Bilim. Araştırmaları Derg.*, vol. 8, no. 3, pp. 550–570, 2023, doi: 10.25307/jssr.1222397.

[9] A. Moelyadi, "Economic Social Cultural Analysis of Sports Coaching Program in East Nusa Tenggara Province," *J. Educ. Heal. Sport*, vol. 10, no. 1, pp. 56–69, 2020, doi: 10.12775/jehs.2020.10.01.007.

[10] H. T. Suppiah, R. Swinbourne, J. Wee, V. Tay, and P. Gastin, "Sleep Characteristics of Elite Youth Athletes: A Clustering Approach to Optimize Sleep Support Strategies," *Int. J. Sports Physiol. Perform.*, vol. 16, no. 9, pp. 1225–1233, 2021, doi: 10.1123/IJSPP.2020-0675.

[11] S. Gaudet, M. Begon, and J. Tremblay, "Cluster analysis using physical performance and self-report measures to identify shoulder injury in overhead female athletes," *J. Sci. Med. Sport*, vol. 22, no. 3, pp. 269–274, 2019, doi: 10.1016/j.jsams.2018.09.224.

[12] I. Masruroh, R. P. Sari, M. P. Pamungkas, and I. Ayuwanti, "Cluster Analysis for School Grouping Based on the Number of Students in East Lampung Regency Using the Ward Method," *Inf dan Mat.*, vol. 2, no. 1, pp. 46–57, 2023.

[13] A. Saladine and E. Prastyansyach, "Analysis of Province Grouping in Indonesia Based on Performance Indicators of the Cattle Farming Sector in 2022," *Zoologi.*, vol. 3, no. 1, 2025, doi: 10.62951/zoologi.v3il.93.

[14] D. Widyadhana, R. B. Hastuti, I. Kharisudin, and F. Fauzi, "Comparison of K-Means and Average Linkage Cluster Analysis for Poverty Clustering in Central Java Province," *Prism. Pros. Semin. Nas. Mat.*, vol. 4, no. 2, pp. 584–594, 2021.

[15] S. Miller, "An analysis of the rate of academic clustering and the types of majors chosen by divisions I, II, and III intercollegiate athletes," *J. Study Sport. Athletes Educ.*, vol. 16, no. 2, pp. 97–113, 2022, doi: 10.1080/19357397.2021.1916305.

[16] L. Yingying, S. Chiusano, and I. V. D'Elia, "Modeling athletic performance Using clustering techniques," *Proc. Third Int. Symp. Electron. Commer. Secur. Work. '10)*, pp. 169–171, 2010.

[17] J. W. Teunissen *et al.*, "'There is more that unites us than divides us'. Optimizing talent transfer processes by clustering 34 sports by their task, individual and environmental similarities," *Front. Sport. Act. Living*, vol. 6, pp. 1–9, 2024, doi: 10.3389/fspor.2024.1445510.

[18] C. Molan, R. Arnold, S. Kelly, E. Toomey, and J. Matthews, "An exploration of performance management processes used within Olympic sport programmes," *J. Appl. Sport Psychol.*, vol. 34, no. 4, pp. 713–733, 2022, doi: 10.1080/10413200.2021.1894506.

[19] W. M. Charmant, P. J. van der Wees, J. B. Staal, R. van Cingel, J. M. Sieben, and R. A. de Bie, "A framework exploring the therapeutic alliance between elite athletes and physiotherapists: a qualitative study," *BMC Sports Sci. Med. Rehabil.*, vol. 13, no. 1, pp. 1–13, 2021, doi: 10.1186/s13102-021-00348-3.

[20] T. McGarry, J. Perl, and M. Lames, "Self-organizing maps and cluster analysis in elite and sub-elite athletic performance," *Complex Syst. Sport*, pp. 208–226, 2013, doi: 10.4324/9780203134610.

[21] M. Grasdalsmoen, B. Clarsen, and B. Sivertsen, "Mental Health in Elite Student Athletes: Exploring the Link Between Training Volume and Mental Health Problems in Norwegian College and University Students," *Front. Sport. Act. Living*, vol. 4, pp. 1–12, 2022, doi: 10.3389/fspor.2022.817757.

[22] L. M. Boleng *et al.*, "Development of norms of physical ability tests for students aged 13-18 years," *Cakrawala Pendidik.*, vol. 42, no. 2, pp. 415–432, 2023, doi: 10.21831/cp.v42i2.55493.

[23] J. J. Fountain and P. S. Finley, "Academic Clustering: A Longitudinal Analysis of a Division I Football Program," *J. Issues Intercoll. Athl.*, vol. 4, pp. 24–41, 2011.

[24] S. Ita, "Global Conferences Series : Determining Dominant Physical Factors in Takraw Service Capability," *Glob. Conf. Ser. Soc. Sci. Educ. Humanit.*, vol. 5, pp. 148–153, 2020, doi: 10.32698/GCS-PSSHERS363.

[25] A. C. T. Smith and C. Stavros, "Exploring the Progressive Use of Performance Enhancing Substances by High-Performance Athletes," *Subst. Use Misuse*, vol. 55, no. 6, pp. 914–927, 2020, doi: 10.1080/10826084.2019.1711412.

[26] A. Kuettel, A. K. Pedersen, and C. H. Larsen, *To Flourish or Languish, that is the question: Exploring the mental health profiles of Danish elite athletes*, vol. 52. 2021. doi: 10.1016/j.psychsport.2020.101837.

[27] K. H. Abdullah, "A Bibliometric and Scoping Retrospective of Sepak Takraw," *Asian J. Arts, Cult. Tour.*, vol. 4, no. 4, pp. 23–37, 2023, doi: 10.55057/ajact.2022.4.4.3.

[28] B. Jiang *et al.*, "Data Analysis of Soccer Athletes' Physical Fitness Test Based on Multi-View Clustering," *J. Phys. Conf. Ser.*, vol. 1060, no. 1, pp. 1–6, 2018, doi: 10.1088/1742-6596/1060/1/012024.

[29] R. M. Ryan and E. L. Deci, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.," *American Psychologist*, vol. 55, no. 1. American Psychological Association, US, pp. 68–78, 2000. doi: 10.1037/0003-066X.55.1.68.

[30] J. A. Moreno-Murcia, E. H. Hernández, L. C. Marín, and J. L. Nuñez, "Coaches' motivational style and athletes' fear of failure," *Int. J. Environ. Res. Public Health*, vol. 16, no. 9, 2019, doi: 10.3390/ijerph16091563.

[31] N. Martínez-González, F. L. Atienza-González, L. González-García, and I. Balaguer, "The motivational climate perceived by young soccer players regarding their coaches, parents, and peers on sport optimal functioning: a cluster analysis," *Front. Psychol.*, vol. 16, no. July, pp. 1–13, 2025, doi: 10.3389/fpsyg.2025.1564391.

[32] E. L. Deci, R. M. Ryan, R. J. Vallerand, and L. G. Pelletier, "Motivation and Education: The Self-Determination Perspective," *Educ. Psychol.*, vol. 26, no. 3–4, pp. 325–346, 1991, doi: 10.1080/00461520.1991.9653137.

[33] G. L. M. Freire, A. R. Contreira, J. F. V. N. DE MORAES, D. V. DE OLIVEIRA, L. Fiorese, and J. R. A. Do Nascimento Junior, "Coach-athlete relationship, team cohesion, and motivation in Brazilian youth athletes: a cluster analysis," *Hum. Mov.*, vol. 24, no. 3, pp. 44–53, 2023, doi: 10.5114/hm.2023.116531.

[34] A. Mahmudan, "Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering," *J.*

Mat. Stat. dan Komputasi, vol. 17, no. 1, pp. 1–13, 2020, doi: 10.20956/jmsk.v17i1.10727.

[35] E. Marheni, E. Purnomo, and F. I. Cahyani, "The Role of Motivation in Increasing Achievement : Perspective Sports Psychology," *Adv. Heal. Sci. Res. (AHSR), Vol. 7 2nd Int. Conf. Sport. Sci. Heal. 2018 (ICSSH 2018)*, vol. 7, pp. 59–62, 2018, doi: 10.2991/icssh-18.2019.14.

[36] F. Claver, L. M. Martínez-Aranda, M. Conejero, and A. Gil-Arias, "Motivation, Discipline, and Academic Performance in Physical Education: A Holistic Approach From Achievement Goal and Self-Determination Theories," *Front. Psychol.*, vol. 11, pp. 1–11, 2020, doi: 10.3389/fpsyg.2020.01808.

[37] R. E. Reigal, A. V Juan, J. P. Morillo-baro, and A. Hern, "Psychological Profile, Competitive Anxiety, Moods and Self-E ffi cacy in Beach Handball Players," *Int. J. Adv. Sport Sci. Res. J. Environ. Res. Public Heal.*, vol. 17, no. 241, pp. 1–13, 2020, doi: doi:10.3390/ijerph17010241.

[38] C. D. McLaren and K. S. Spink, "Team Member Communication and Perceived Cohesion in Youth Soccer," *Commun. Sport*, vol. 6, no. 1, pp. 111–125, 2018, doi: 10.1177/2167479516679412.

[39] Y. Krasmik *et al.*, "Motivational determinants of athletes' self-realisation depending on their professional qualification," *BMC Psychol.*, vol. 12, no. 1, pp. 1–16, 2024, doi: 10.1186/s40359-024-01895-3.

[40] M. Manninen, R. Dishman, Y. Hwang, E. Magrum, Y. Deng, and S. Yli-Piipari, "Self-determination theory based instructional interventions and motivational regulations in organized physical activity: A systematic review and multivariate meta-analysis," *Psychol. Sport Exerc.*, vol. 62, pp. 1–17, 2022, doi: 10.1016/j.psychsport.2022.102248.

[41] M. Ayranci and M. K. Aydin, "The complex interplay between psychological factors and sports performance: A systematic review and meta-analysis," *PLoS One*, vol. 20, pp. 1–24, 2025, doi: 10.1371/journal.pone.0330862.

[42] M. G. Trotter, T. J. Coulter, P. A. Davis, D. R. Poulus, and R. Polman, "Social Support, Self-Regulation, and Psychological Skill Use in E-Athletes," *Front. Psychol.*, vol. 12, pp. 1–10, 2021, doi: 10.3389/fpsyg.2021.722030.

[43] A. J. Clements and C. Kamau, "Understanding students' motivation towards proactive career behaviours through goal-setting theory and the job demands–resources model," *Stud. High. Educ.*, vol. 43, no. 12, pp. 2279–2293, 2018, doi: 10.1080/03075079.2017.1326022.

[44] A. Chaeroni *et al.*, "Improving Soccer Coaching: Considerations of Individual Learning Styles, Intelligence Levels, and Motivation," *Retos*, vol. 58, pp. 377–383, 2024, doi: 10.47197/retos.v58.104086.

[45] N. B. Feddersen, M. A. B. Keis, and A. M. Elbe, "Coaches' perceived pitfalls in delivering psychological skills training to high-level youth athletes in fencing and football," *Int. J. Sport. Sci. Coach.*, vol. 16, no. 2, pp. 249–261, 2021, doi: 10.1177/1747954120959524.

[46] J. E. Bissett, E. Kroshus, and S. Hebard, "Determining the role of sport coaches in promoting athlete mental health: A narrative review and Delphi approach," *BMJ Open Sport Exerc. Med.*, vol. 6, no. 1, pp. 1–9, 2020, doi: 10.1136/bmjsem-2019-000676.

[47] L. Davis and S. Jowett, "Coach-athlete attachment and the quality of the coach-athlete relationship: implications for athlete's well-being," *J. Sports Sci.*, vol. 32, no. 15, pp. 1454–1464, 2014, doi: 10.1080/02640414.2014.898183.

[48] H. Jin, S. Kim, A. Love, Y. Jin, and J. Zhao, "Effects of leadership style on coach-athlete relationship, athletes' motivations, and athlete satisfaction," *Front. Psychol.*, vol. 13, pp. 1–14, 2022, doi: 10.3389/fpsyg.2022.1012953.

[49] S. Jowett, J. W. Adie, K. J. Bartholomew, S. X. Yang, H. Gustafsson, and A. Lopez-Jiménez, "Motivational processes in the coach-athlete relationship: A multi-cultural self-determination approach," *Psychol. Sport Exerc.*, vol. 32, pp. 143–152, 2017, doi: 10.1016/j.psychsport.2017.06.004.

[50] V. Roşca, "The coach-athlete communication process: Towards a better human resources management in sport," *Manag. Res. Pract.*, vol. 2, no. 3, pp. 275–283, 2010.

[51] T. Larsen *et al.*, "Creating a supportive environment among youth football players: A qualitative study of French and Norwegian youth grassroots football coaches," *Health Educ.*, vol. 115, no. 6, pp. 570–586, 2015, doi: 10.1108/HE-04-2014-0054.

[52] P. Yin, M. F. Bin Abdullah, and M. F. B. H. Hassan, "The key factors of talent development environment for campus football athletes in China," *Cogent Soc. Sci.*, vol. 10, no. 1, pp. 1–14, 2024, doi: 10.1080/23311886.2024.2431590.

[53] S. Afrizal *et al.*, "Integration of Life Skills in Football Training Programs in the Context of Positive Youth Development," *Int. J. Disabil. Sport. Heal. Sci.*, vol. 7, no. 1, pp. 29–36, 2024, doi: 10.33438/ijdshs.1368983.

[54] S. Super, K. Verkooijen, and M. Koelen, "The role of community sports coaches in creating optimal social conditions for life skill development and transferability–a salutogenic perspective," *Sport. Educ. Soc.*, vol. 23, no. 2, pp. 173–185, 2018, doi: 10.1080/13573322.2016.1145109.

[55] J. Vaughan, C. J. Mallett, P. Potrac, M. A. López-Felip, and K. Davids, "Football, Culture, Skill Development and Sport Coaching: Extending Ecological Approaches in Athlete Development Using the Skilled Intentionality Framework," *Front. Psychol.*, vol. 12, pp. 1–13, 2021, doi: 10.3389/fpsyg.2021.635420.

[56] J. Bae, T. Lim, and D. M. O'Sullivan, "The role of sport coaches on student-athletes' life skills development and transfer: Using multilevel model analysis," *Int. J. Sport. Sci. Coach.*, vol. 19, no. 3, pp. 944–955, 2024, doi: 10.1177/17479541231217368.