International Journal of Sports Science 2025, 15(2): 47-57
DOI: 10.5923/j.sports.20251502.03

A Mathematical Model Based on Pairwise Comparison and Genetic Algorithms to Assist Tennis Coaches of Non-professional Tennis Players

J. Werneck Allen[1], Antonio Juarez S. M. de Alencar[1,*], William A. Jr.[1], Elisangela L. Faria[1,2], Marcelo P. Albuquerque[2], Clecio R. Bom[2], Paulo J. Russano[2]

[1]ImagineLabs Sports, Rio de Janeiro, Brazil
[2]Brazilian Center for Physics Research (CBPF), Rio de Janeiro, Brazil

Abstract This paper introduces a mathematical model that allows non-professional tennis players to be classified. As a result, it evaluates the shortest way to improve their performance. The model is based on expert opinion and genetic algorithms. A large variety of performance indicators are considered by the model. These indicators can be easily extended to accommodate the needs of players and their respective coaches. Expert opinion is captured using the pairwise comparison ideas that served as the basis for the development of Saaty's Analytic Hierarchy Process (AHP). The model provides valuable insights to both coaches and tennis players.

Keywords Pairwise comparison, AHP, Genetic algorithms, Tennis players, Tennis coaches, Coach assistance

1. Introduction

Throughout history, statistical analysis has been an important part of sports activities. Over time, amateur practitioners, professional players, coaches, and fans have relied on the analysis of performance indicators to identify the strong and weak points of individual athletes, the teams to which they belong, and their adversaries [1].

The methods used in these analyses vary considerably. They include basic statistical calculations such as frequency, mean, standard deviation, and correlation [2], and more advanced methods, such as logistic regression [3], analysis of variance [4] and resampling [5].

Over the last few decades, sport analytics has seen the addition of new methods and techniques coming from the decision-making research area, such as the Analytic Hierarchy Process [6], the Analytic Network Process [7], TOPSIS [8], and VIKOR [9]. Moreover, mathematics and computer science have also made their contribution to sports analytics with the introduction of methods such as fuzzy logic [10], random forest [11] and neural networks [12]. Dindorf *et al.* present a comprehensive discussion on the use of artificial intelligence in sports [13].

Nevertheless, in recent years, new developments in computer vision and biometric wearables have allowed for a whole new set of performance indicators to be captured and analyzed in real time. By analyzing images captured by a camera, one can figure out an athlete's reaction time, strength, ground speed, positioning, and strategy, among other skill indicators [14]. In addition, biometric wearables can be used to monitor an athlete's vital signs, such as heart rate, blood pressure, stress level, and oxygen saturation [15].

All of this has served as a basis to build computer applications that can capture and process information on the physical, tactical, and technical aspects of sports practice while these activities are taking place or immediately after. The processed information can be composed of general statistics of an athlete's skills, the rank of athletes in accordance with their performance in sports events, and predictions about the results of tournaments. The Internet site Slashdot presents an extensive list of sports computer applications that are available on the market [16].

This paper builds upon the current state of these computer applications. It presents a mathematical model based on the use of the pairwise comparison technique used by Saaty when building the Analytic Hierarchy Process (AHP) [17]. The model uses the expert opinion of both coaches and professional tennis players to analyze the performance indicators frequently captured by tennis-monitoring computer applications.

These indicators are then used to rate athletes' skills and classify them into distinct categories. Only the athletes' personal skills are considered by the classification process, i.e., it does not take into account the athletes' performance in tournaments. This allows non-professional tennis players

* Corresponding author:
ajsmalencar@gmail.com (Antonio Juarez S. M. de Alencar)
Received: Oct. 9, 2025; Accepted: Nov. 3, 2025; Published: Nov. 14, 2025
Published online at http://journal.sapub.org/sports

to be compared fairly to other players without the need to play tennis competitively. The shortest course of action for an athlete to ascend in the category hierarchy is provided with the support of genetic algorithms.

2. Background

2.1. Pairwise Comparison

Given a set of entities $E = \{e_1, e_2, \cdots, e_n\}$ and a criterion C, pairwise comparison is a preference prioritization process in which all the objects in E are compared among themselves in respect to C. In the pairwise comparison process, two entities are compared at a time. These comparisons may be represented as an $n \times n$ matrix, where C_{e_i,e_j} is the result of comparing entity e_i with entity e_j, for $1 \leq i, j \leq n$. See Table 1 in this respect.

Table 1. The Pairwise Comparison Matrix (PCM)

Criterion $\boldsymbol{C}$	$\boldsymbol{e_1}$	$\boldsymbol{e_2}$	$\cdots$	$\boldsymbol{e_n}$
$\boldsymbol{e_1}$	C_{e_1,e_1}	C_{e_1,e_2}	$\cdots$	C_{e_1,e_n}
$\boldsymbol{e_2}$	C_{e_2,e_1}	C_{e_2,e_2}	$\cdots$	C_{e_2,e_n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{e_n}$	C_{e_n,e_1}	C_{e_n,e_2}	$\cdots$	C_{e_n,e_n}

If the comparison criterion C is consistent, then C_{e_i,e_j} yields a comparison result that is the reciprocal of the result yielded by C_{e_j,e_i}. For example, if the criterion in place is height and Peter is much taller than John, then John is much shorter than Peter. It should be noted that C_{e_i,e_i} is the result of comparing an object with itself and, consequently, the comparison result and its reciprocal are the same. For instance, in respect to height, Peter is the same height as himself.

The use of pairwise comparison in science is due to the work of the psychometrician L. L. Thurstone at the beginning of the 20th century [18]. An introduction to pairwise comparison is presented in [19]. Recent Advances in pairwise comparison can be found in [20].

2.2. The Analytic Hierarchy Process (AHP)

The AHP is a decision-making method that has been conceived to derive ratio scales from the elements of a given set of entities by making extensive use of pairwise comparison [17]. The ratio scales reflect the relative strength of preferences of one entity over another. The entities being compared may be evaluated by criteria based on abstract perceptions of reality, such as beauty, comfort and safety, and of concrete nature, such as height, weight, and speed.

In the AHP, several criteria may be combined to rank the preferences among entities. To facilitate reasoning and avoid inconsistencies, such a combination of criteria is organized in a hierarchy. At the top of the hierarchy, one finds the overall objective to be reached. For example, to select the best player amongst a group of tennis players.

The intermediate levels consist of the criteria and sub-criteria to be considered when ranking preferences among entities. For instance, the skills of tennis players may be compared using criteria such as technical skill, physical capacity, and tactical ability. Finally, at the bottom of the hierarchy reside the entities to be ranked. For example, the actual tennis players. In the AHP jargon, these entities are called "alternatives".

In the AHP, the result of comparing entities is evaluated with the support of a comparison scale. Table 2 shows a simplified form of this scale, as proposed by Saaty in [17].

Table 2. Scale of Comparison

Scale	Definition in Respect to the Criterion Being Considered
1	The entities being compared are equally relevant
3	One entity being compared is slightly more relevant than the other
5	One entity being compared is moderately more relevant than the other
7	One entity being compared is considerably more relevant than the other
9	One entity being compared is extremely more relevant than the other
2, 4, 6, 8	Intermediate values may be used when needed.

Moreover, the result of comparing entities using a criterion C is represented as the $n \times n$ matrix shown on the left part of Table 1. It should be noted that because the comparison scale presented in Table 2 is used, the elements on the diagonal of the $n \times n$ matrix are all equal to 1, i.e. C_{o_i,o_i} is 1.

2.3. An Example

For example, consider that a group of non-professional tennis players is to be ranked according to their tactical ability, technical skill, and physical capacity. Table 3 presents an AHP comparison matrix used to determine the relative importance of these indicators to play tennis. In general, matrices like this are based on the opinion of a group of coaches and professional tennis players. See the acknowledgements part of this article in this regard.

Table 3. Comparing Common Skills Used to Rank Non-Professional Tennis Players

	Tactical ability	Technical skill	Physical Capacity		Weights
Tactical ability	**1**	1/3	1/7		0.093
Technical skill	3	**1**	1/2	⇒	0.292
Physical Capacity	7	2	**1**		0.615
				Total	1.000

Table 4. Random Consistency Index

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

It should be noted that according to the opinion of those experts, to reach the overall objective of the comparison process:

- Technical skill is rated 3 when compared with tactical ability, indicating that they are slightly more relevant than the latter.
- Physical capacity is rated 7 when compared with tactical ability; thus, the former is considerably more relevant than the latter.
- Physical capacity is rated 2 when compared with technical skill, hence it is marginally more relevant than the latter.

According to Saaty, the result of comparing entities using a criterion C is the normalized eigenvector associated with the highest eigenvalue of the $n \times n$ matrix, i.e. L_{max}. In the example presented in Table 3, the vector "Weights" is this eigenvector. It should be noted that it expresses the relative importance of those skills, abilities and capacity in the form of numerical values. The higher the numerical value, the more important the indicator is.

Because inconsistencies may arise during the comparison process, Saaty proposes a consistency rate $= CI/RI$, where CI, the consistency index is given by

$$CI = (L_{max} - n)/(n - 1)$$

and RI, the random consistency index, is obtained from Table 4.

It is required that $CR < 10\%$, otherwise the decision makers must revise their decisions [21]. The CI, RI and CR of the comparison matrix presented in Table 3 are respectively 0.002, 0.58 and 0.32%. Therefore, no revisions of previously made decisions are necessary

Over the years, an extensive bibliography has been developed around the Analytic Hierarchy Process, and many articles and books have been published in the scientific and business areas [22]. An introduction to the AHP can be found in [23,24].

2.4. The AHP and Pairwise Comparison in the Practice of Tennis

Both the AHP and pairwise comparison are not strangers to the art of tennis. Over time, they have been used extensively to rank and pinpoint promising tennis players and make predictions about the outcome of tennis matches.

For example, Bozóki, Csató, and Temesi [25] use incomplete pairwise comparison matrices to rank professional tennis players who have been at the top of the Association of Tennis Professionals (ATP) rank over the last 40 years. In their work, the eigenvector [17] and the logarithm least squares method [26] are used to figure out the weights of the incomplete matrices. An important aspect of Bozóki, Csató and Temesi's work is that it does not require that all players have faced each other on the court. In [27], Temesi, Szádoczki and Bozóki adopt a similar approach to rank top women tennis players.

Garcia and Mori apply pairwise comparisons to help identify the greatest tennis players of all time or GOAT [28]. Data from the Association of Tennis Professionals (ATP) and the Women's Tennis Association (WTA) is used to select the male and female greatest players. Not surprisingly, well-known players such as Serena Williams, Steffi Graf, and Martina Navratilova come at the top of the women GOAT's list and Novak Djokovic, Roger Federer, and Rafael Nadal at the top of the male list.

In [29] Wu employs the AHP and statistical tests [30] to quantify momentum in tennis matches. His work shows that, in mathematical terms, momentum can be defined as the probability of scoring or losing a point as a function of the outcome of the preceding point. In addition, Wu's findings refute the long-held belief that performance fluctuations in tennis matches are random events. As a result, it is up to coaches to better prepare their athletes to create and maintain favorable momentum.

The AHP, XGBoost [31] and genetic algorithms [32] are utilized by Kang *et al.* to enhance tennis match predictions [33]. Their article uses the AHP to identify key factors influencing momentum in matches. Subsequently, the weights of these factors are optimized using XGBoost and genetic algorithms with the goal of improving the accuracy of predictions.

Hraste, Đurovié and Stanišié rely on the AHP to explain the decision-making priorities of offensive tennis players when they are playing offensively and defensively. The relative importance of these priorities is based on the opinion of seven coaches who stand out on the world circuit. By being aware of these decision-making priorities, coaches can better prepare their athletes for the challenges of tennis matches and tournaments [34].

Zang *et al.* [35] combine AHP with the entropy weighted method [36], TOPSIS [8], and principal component analysis [37] to predict tennis players' scoring ability based on their performance during a match in real time. As a result, the prediction accuracy of women's tennis matches surpassed 0.95 in major tournaments.

Lei, Lin and Cao Jr. bring together the AHP and support vector machine (SVM) [38] to predict the outcome of professional tennis matches. The resulting model was tested using data from the Wimbledon Championship 2023. Their model has reached an accuracy of over 0.83 when predicting the winner of male professional matches. Furthermore, the authors claim that their model is more accurate than existing models in predicting the winner of a point [39].

Regardless of having the largest population in the world, China lacks world-class professional tennis players [40]. Guangfu and Kanchanathaweekul have devised a way to help tackle this problem. In [41], they explore the use of the AHP to identify talented young tennis players in the country. By proving the means to distinguish these athletes from the regular players, they expect talented tennis players to receive the necessary support from the Chinese government. As a result, the presence of Chinese athletes in the international circuit may increase substantially.

Despite the vast literature that has connected pairwise comparison and the AHP to the art of tennis, very little has been done to suggest what actions non-professional players should undertake to improve their performance. This article goes towards filling this gap.

2.5. Genetic Algorithms

Genetic algorithms are a family of mathematical optimization methods inspired by the selection process widely observed in nature. The optimization begins by randomly creating an initial population of individuals in the potential solution space. It proceeds by refining this population interactively using selection, crossover, and mutation processes. In each interaction, a population of better-fit individuals is created until a satisfactory solution is found.

In this context, a fitness function is used to indicate how well an individual performs with respect to the optimization objective. By using the fitness function, the selection process chooses the fittest individuals in the population to parent the next generation. In general, this process is probabilistic, where the best-fit individuals have a higher chance of being selected.

In genetic algorithms, the characteristics of individuals who are the potential solution to the problem at hand are often encoded as a sequence of 0s and 1s. These sequences are frequently referred to as the individuals' chromosomes.

In the crossover process, the chromosomes of two individuals are combined to produce an offspring. This simulates the natural process of reproduction, where the good characteristics of two individuals are combined to yield a better-suited offspring.

The mutation action allows small changes to the characteristics of individuals to be introduced randomly. This allows new areas of the solution space to be considered while helping to prevent the optimization process from being trapped in a local optimal solution.

Although genetic algorithms are the result of contributions of several researchers over time, John H. Holland is credited with the development of this family of optimization methods in the late 1970s and for its popularization [49]. An introduction to Genetic Algorithms can be found in [48].

3. The Model

The following steps have guided the development of the models proposed in this article:

1. Assembling experts in the art of tennis and coaching.
2. Gathering the data and the metadata.
3. Identifying the criteria and sub-criteria used to build the classification model.
4. Defining the scale to evaluate each sub-criterion and the set of alternatives.
5. Building the model.
6. Presenting suggestions on what actions non-professional tennis players should take to improve their performance.

3.1. Assembling Experts

A group of experts was assembled to support the development of the model presented in this article. These experts were selected from experienced tennis players with relevant national and international experience. All of them are involved with tennis in the State of Rio de Janeiro, Brazil, including activities that help others to improve their technical tennis skills and tactical abilities.

This group of experts was asked to help develop a mathematical model capable of classifying non-professional tennis players into categories. Furthermore, this model should be able to provide players with a clear path on how to improve their skills, abilities, and physical capacity step by step. These experts are listed in the acknowledgements section of this article.

3.2. Gathering the Data and Metadata

ImagineLabs (IL) [42] is a startup company based in the city of São Paulo, Brazil. Over the last 14 months IL has been developing a monitoring and training analysis app for tennis. It is called TotalTennis [43]. At this point, hundreds of thousands of video frames have been captured and analyzed by the app, using computer vision and various artificial intelligence models and machine learning algorithms. As a result, a number of statistics have been gathered and made available to support the writing of this article.

These are comprised of statistics about point, set and game disputes, such as duration, score tracking, number of balls exchanged, the height, trajectory and speed of the ball, variety of shot types, hit point of the ball on the court, distance covered by each player, position of the player on the court, balance of each player, and players' reaction time, among others.

3.3. Identifying the Criteria and Sub-criteria

Based on the data available for analysis and the goal established to build the model introduced in this article (see Item 3.1), the aforementioned group of experts devised the criteria and sub-criteria to be used for pairwise comparison. These criteria are shown in Table 5, Table 6 and Table 7.

Even in training, it is well known that it takes two players to play a game of tennis, i.e. the player of interest and his or her adversary. This designation is obviously arbitrary. Any of the players can be designated as the player of interest. Nevertheless, for the purpose of the model developed in this article, the criteria and sub-criteria presented in Table 5,

Table 6 and Table 7 always refer to this player.

It should be observed that each sub-criterion takes a value in a certain numerical interval. Some of these intervals are percentages, while others are not. Nevertheless, to facilitate reasoning, when one of those non-percentage criteria takes an actual value, it is transformed into a percentage value. Let v_i be one of those values. In this case p_i, the corresponding percentage value is given by v_i/UB_i, where UB_i is the upper bound of sub-criterion i. Furthermore, if $v_i < LB_i$, the lower bound of criterion i, then v_i assume the value of LB_i. In addition, if $v_i > UB_i$, then v_i assume the value of UB_i.

It is important to note that for most of the criteria presented in Table 5, Table 6 and Table 7, the closer a value is to their respective upper bound, the better. However, this does not apply to sub-criteria 12, 13, and 14, the average return depth, average serve depth, and average groundstroke depth, respectively.

On a tennis court, the white line that runs parallel to the net at the rear boundary of both sides is called the baseline. Similarly, the white line that runs horizontally close to the net is called the service line, which is delimited by two sidelines forming a rectangle, called the service box. There are two service boxes on each side of the tennis court.

During an exchange of balls between players (i.e., a rally), placing the ball close to the baseline is generally considered a strategic play. It forces one's opponent farther back, reducing his or her offensive options. As a result, this helps to gain control over the rally and increase the probability of scoring points. The same line of reasoning can be applied during service. The closer a ball bounces to any of the lines delimiting the service box, the better. For balls bouncing inside the service box, sub-criterion 12 is the average distance from the bouncing point to the nearest line of the box.

In tennis jargon, 0 meters to the baseline or the lines of the service box indicates that a ball hit one of those lines. If the distance is greater than 0 and smaller than the sub-criteria upper bound, it bounced inside the court limits; otherwise, it is out.

Therefore, for sub-criteria 12, 13, and 14, the closer a value is to their lower bound, the better. As a result, to simplify matters, with respect to these sub-criteria, p_i is calculated as $1-(v_i/UB_i)$. As a result, for all sub-criteria introduced in Table 5, Table 6 and Table 7, the closer their values are to their upper bound the better.

3.4. Defining the Weights Used to Evaluate Each Sub-Criterion

The relative importance of the tactical ability, technical skill and physical capacity criteria is presented in Table 1, see column "Weights" in this respect. The relative importance of the corresponding sub-criteria are presented in Table 8, Table 9, and Table 10.

3.5. Defining the Weights Used to Evaluate the Set of Alternatives

As indicated in Item 3.3, the alternatives' rates take value in the [0, 1] interval. Following the ideas of [44,45,46] on the use of absolute measurement in the AHP, the alternatives' rates are grouped into categories. Column "Categories' Intervals" of Table 11 shows these categories. They were determined based on the consensual opinion of the experts cited in Item 3.1. Those categories are considered effective in distinguishing between low-performing tennis players and those with intermediate or advanced skill levels.

Next, the weights of the categorized rates are obtained for each sub-criterion using pairwise comparison and the AHP. It was found that the weights attributed to those categories follow a near-linear scale.

There is no clear evidence that each individual sub-criterion would necessarily require a different set of weights. As a result, Occam's simplicity principle is used [47] and the same weights are used to rate all the alternatives with respect to the same set of sub-criteria. Column "Categories' Weights" of Table 11 introduces these weights.

3.6. Building the Model

Let Tac, Tec and PC be the weights attributed to tactical ability, technical skill, and physical capacity criteria, respectively, as presented in Table 3. Moreover, allow

- Tac_i be the weight attributed to the i-th tactical ability sub-criterion,
- Tec_j be the weight associated to the j-th technical skill sub-criterion and
- PC_k be the weight given to the k-th physical capacity sub-criterion,

as introduced in Table 8, Table 9 and Table 10.

Furthermore, let

- $0 \le p_{Tec_j}$, p_{Tac_i} and $p_{PC_k} \le 1$ be the percentage scores given to a player P of interest as calculated by a monitoring and analysis app for tennis, for example, the "TotalTennis" app, and
- $wcat_{Tac}(p_{Tac_i})$, $wcat_{Tec}(p_{Tec_j})$ and $wcat_{PC}(p_{PC_K})$ be the weights calculated for the category and sub-criterion into which p_{Tac_i}, p_{Tec_j} and p_{PC_K} are grouped. See Table 11 in this regard.

In these circumstances, the total score attributed to a tennis player P, i.e. $sc(P)$, is given by

$$sc(P) = Tac \times \left(\sum Tac_i \times wcat_{Tac}(p_{Tac_i})\right) + Tec \times \left(\sum Tec_j \times wcat_{Tec}(p_{Tec_j})\right) + PC \times \left(\sum PC_k \times wcat_{PC}(p_{PC_K})\right)$$

Table 5. Physical capacity criteria and sub-criteria to be considered for pairwise comparison

Physical Capacity			
Item	**Sub-criteria (performance indicators)**	**Lower Bound**	**Upper Bound**
1	*Average Ball Groundstroke Speed* - average speed of baseline strokes (forehands, backhands, and slices only), measured from the moment of ball contact to the opponent's contact, second bounce, or error (into the net or out), when applicable	35 km/h	130 km/h
2	*Average Service Speed* - average speed of all serves combined (including both first and second serves)	35 km/h	190 km/h
3	*Average Player Displacement Speed* - average speed of the player's displacement while the ball is in play	0.1 m/s	5 m/s
4	*Rallies per Point* – average number of strokes per point (total strokes / total points)	1 ball	25 balls
5	*% Balanced Shots* – percentage of strokes executed with proper body balance, relative to the total number of strokes.	0%	100%

Table 6. Tactical ability criteria and sub-criteria to be considered for pair-wise comparison

Tactical Ability			
Item	**Sub-criteria (performance indicators)**	**Lower Bound**	**Upper Bound**
6	*% Points Won on Serve* – number of points won while serving divided by the total number of service points played	0%	100%
7	*% Points Won on Return* – number of points won while returning, divided by the total number of return points played	0%	100%
8	*% Ideal Court Positioning* – percentage of shots in which the player was in the ideal court position at the moment of the opponent's shot. The ideal position is defined by the bisector of the potential return angle	0%	100%
9	*% Ideal Shot Selection* – ability to choose tactically appropriate shots based on timing and court position during the rally, compared to common decision-making patterns	0%	100%
10	*% Stroke Consistency* – accuracy percentage for each type of stroke, based on the number of successful shots relative to the total executed.	0%	100%
11	*% Serve Consistency* – percentage of serves that landed in, relative to the total number of serves attempted.	0%	100%

Table 7. Technical skill criteria and sub-criteria to be considered for pair-wise comparison

Technical Skill			
Item	**Sub-criteria (performance indicators)**	**Lower Bound**	**Upper Bound**
12	*Average Return Depth* – average distance of return bounces measured from the opponent's baseline.	0 m	11.885 m
13	*Average Serve Depth* – average distance from the bouncing point of the serve to the nearest service line (center or sideline), whichever is closer.	0 m	6,4 m
14	*Average Groundstroke Depth* – average distance of groundstroke bounces (forehands, backhands, and slices) measured from the opponent's baseline	0 m	11.885 m
15	*% Control of Shot Direction and Depth* – measurement of variation and precision in shot placement based on bounce zones defined in the app, accounting for both direction and depth, calculated from the point of contact to the bounce location.	0%	100%
16	*% Proper Technique Execution* – percentage of strokes executed with technically sound form relative to the number of total strokes when compared to a database of technically correct executions.	0%	100%

3.7. An Example

To simplify matters, consider that all $p_{Tac_i} = 0.48$, $p_{Tec_j} = 0.55$ and $p_{PC_K} = 0.61$. Therefore, in these circumstances:

$$\sum Tac_i \times p_{Tac_i} = 0.037 \times wcat_{Tac}(0.48) + 0.057 \times wcat_{Tac}(0.48) + 0.104 \times wcat_{Tac}(0.48) + 0.180 \times wcat_{Tac}(0.48) + 0.216 \times wcat_{Tac}(0.48) + 0.406 \times wcat_{Tac}(0.48) = 0.037 \times 0.087 + 0.057 \times 0.087 + 0.104 \times 0.087 + 0.180 \times 0.087 + 0.216 \times 0.087 + 0.406 \times 0.087 = 0.087$$

$$\sum Tec_j \times p_{Tec_j} = 0.041 \times wcat_{Tec}(0.55) + 0.080 \times wcat_{Tec}(0.55) + 0.230 \times wcat_{Tec}(0.55) + 0.230 \times wcat_{Tec}(0.55) + 0.419 \times wcat_{Tec}(0.55) = 0.041 \times 0.109 + 0.080 \times 0.109 + 0.230 \times 0.109 + 0.230 \times 0.109 + 0.419 \times 0.109 = 0.109$$

Table 8. The weights of the physical capacity sub-criteria (CR = 6.72%)

Physical capacity sub-criteria	Avg Serve Speed (Item 2)	Avg Gr Stroke Speed (Item 1)	% Balanced Shots (Item 5)	Avg Rallies per Point (Item 4)	Avg Disp. Speed (Item 3)		Weights
Avg Serve Speed (Item 2)	1	1/3	1/8	1/9	1/9	⇒	0.031
Avg Gr Stroke Speed (Item 1)	3	1	1/6	1/7	1/7		0.056
% Balanced Shots (Item 5)	8	6	1	1/2	1/2		0.220
Avg Rallies per Point (Item 4)	9	7	2	1	1		0.346
Avg Disp. Speed (Item 3)	9	7	2	1	1		0.346
					Total		1.000

$$\begin{aligned}\sum PC_k \times p_{PC_K} = {} & 0.031 \times wcat_{PC}\,(0.61) + 0.056 \times wcat_{PC}\,(0.61) + 0.220 \times wcat_{PC}\,(0.61) \\ & +0.346 \times wcat_{PC}\,(0.61) + 0.346 \times wcat_{PC}\,(0.61) \\ = {} & 0.031 \times 0.130 + 0.056 \times 0.130 + 0.220 \times 0.130 + 0.346 \times 0.130 + 0.346 \times 0.130 = 0.130\end{aligned}$$

As a result,

$$\begin{aligned}sc(P) &= Tac\ \times 0.087 + Tec\ \times 0.109 + \mathrm{PC}\ \times 0.130 \\ &= (0.093 \times 0.087) + (0.292 \times 0.109) + (0.615 \times 0.130) \\ &= 0.120\end{aligned}$$

Table 9. The weights of the tactical ability sub-criteria (CR=1.43%)

Tactical ability sub-criteria	% Points Won on Serve (Item 6)	% Ideal Court Posit (Item 8)	% Serve Consistency (Item 11)	% Points Won on Return (Item 7)	% Ideal Shot Selection (Item 9)	% Stroke Consistency (Item 10)		Weights
% Points Won on Return (Item 7)	1	1/2	1/3	1/5	1/6	1/8	⇒	0.037
% Points Won on Serve (Item 6)	2	1	1/2	1/3	1/5	1/7		0.057
% Serve Consistency (Item 11)	3	2	1	1/2	1/2	1/4		0.104
% Ideal Court Posit (Item 8)	5	3	2	1	1	1/3		0.180
% Ideal Shot Selection (Item 9)	6	5	2	1	1	1/2		0.216
% Stroke Consistency (Item 10)	8	7	4	3	2	1		0.406
						Total		1.000

Table 10. The weights of the technical skill sub-criteria (CR = 1,43%)

Technical skill sub-criteria	Avg Serve Depth (Item 13)	Avg Return Depth (Item 12)	% Ctrl Shot Dir and Depth (Item 15)	Avg Gr Stroke Depth (Item 14)	% Proper Tech Exec (Item 16)		Weights
Avg Serve Depth (Item 13)	1	1/2	1/6	1/6	1/9	⇒	0.041
Avg Return Depth (Item 12)	2	1	1/3	1/3	1/5		0.080
% Proper Tech Exec (Item 16)	6	3	1	1	1/2		0.230
Avg Gr Stroke Depth (Item 14)	6	3	1	1	1/2		0.230
% Ctrl Shot Dir and Depth (Item 15)	9	5	2	2	1		0.419
					Total		1.000

3.8. Defining the Players' Categories

Because $0 \leq sc(P) \leq 1$, every player whose performance is evaluated by the model presented in this article is rated in that interval. As a result, the group of experts (see Item 3.1 in this respect) devised a set of categories in which tennis players are placed according to their skill, ability and capacity. Table 12 shows these categories.

3.9. Presenting Suggestions

To present suggestions on what a tennis player should do to reach the next category, it suffices to resolve the following optimization problem:

$$Min\left(\sum inc_{Tac_i}\right) + \left(\sum inc_{Tec_j}\right) + \left(\sum inc_{PC_k}\right) \quad (1)$$

subject to

- $0 \leq inc_{Tac_i}, inc_{Tec_j}, inc_{PC_K} \leq 1$,
- $cs\left(Tac \times \left(\sum Tac_i \times wcat_{Tac}\left(p_{Tac_i} + inc_{Tac_i}\right)\right) + Tec \times \left(\sum Tec_j \times wcat_{Tec}\left(p_{Tec_j} + inc_{Tec_j}\right)\right) + PC \times \left(\sum PC_k \times wcat_{PC}\left(p_{PC_K} + inc_{PC_k}\right)\right)\right) = LB_{Next\ Level}$,
- and
- $0 \leq p_{Tac_i} + inc_{Tac_i}, p_{Tec_j} + inc_{Tec_j}, p_{PC_K} + inc_{PC_k} \leq 1$,

where

- Tac, Tec and PC are the weights defined in Table 3,
- Tac_i, Tec_j and PC_k are the weights defined in Table 6, Table 7 and Table 8,
- p_{Tac_i}, p_{Tec_j} and p_{PC_k} are the percentage rates attributed to a player of interest,
- $wcat_{Tac}\left(p_{Tac_i}\right)$, $wcat_{Tec}\left(p_{Tec_j}\right)$ and $wcat_{PC}\left(p_{PC_K}\right)$ are the categorical weights attributed to a player of interest as introduced in Table 11,
- $LB_{Next\ Level}$ is the lower bound value of the category immediately higher than that of the player of interest. and
- inc_{Tac_i}, inc_{Tec_j} and inc_{PC_k} are the suggested rate increments that the player of interest should achieve or surpass to enter that category.

This optimization problem can be solved with the use of genetic algorithms [48]. In the example introduced in Item 3.6, the player of interest is rated at 0.57. Therefore, according to Table 12, he or she is an "Intermediate". When the system of equations presented above is applied to the player's performance indicators, it shows that a mere average increase of 3.25% in the performance indicators presented in Table 5 is enough for he or she to move to the next category.

With the intention of avoiding asking too much effort from a tennis play in just one step, the suggested increments inc_{Tac_i}, inc_{Tec_j} and inc_{PC_k} can have their contribution limited with the introduction of the following restrictions:

- $inc_{Tac_i} \leq l\%_{Tac_i} \times p_{Tac_i}$,
- $inc_{Tec_j} \leq l\%_{Tec_j} \times p_{Tec_j}$

and

- $inc_{PC_k} \leq l\%_{PC_k} \times p_{PC_k}$,

where $l\%_{Tac_i}$ $l\%_{Tec_j}$, and $l\%_{PC_k}$ are percentage limits imposed on each of the sets of sub-criteria.

Moreover, these limits can be used to select the kind of suggestions that one is willing to present to the player of interest. For example, if both $l\%_{Tac_i}$ and $l\%_{PC_k}$ are set to zero and $l\%_{Tec_j}$ is set to a number higher than that, the system of equations would be allowed to present only technical skill improvement suggestions. The same line of reasoning can be used to present only physical capacity or tactical ability improvement suggestions.

Table 11. Alternatives Categorial Weights (CR = 0%)

Categories' Intervals		Categories' Weights
[0.0, 0.1]		0.022
(0.1, 0.2]		0.022
(0.2, 0.3]		0.043
(0.3, 0.4]		0.065
(0.4, 0.5]	⇒	0.087
(0.5, 0.6]		0.109
(0.6, 0.7]		0.130
(0.7, 0.8]		0.152
(0.8, 0.9]		0.174
(0.9, 1.0]		0.196
	Total	1,000

Table 12. Tennis players' categories

Interval	Category
$\mathbf{0 \leq sc(P) < 0.3}$	Beginner
$\mathbf{0.3 \leq sc(P) < 0.6}$	Intermediate
$\mathbf{0.6 \leq sc(P) < 0.8}$	Upper Intermediate
$\mathbf{0.8 \leq sc(P) < 0.9}$	Advanced
$\mathbf{0.9 \leq sc(P) \leq 1.0}$	Semi-pro

4. Conclusions

Mobile phone apps such as TotalTennis have opened a whole new door to the monitoring and analysis of tennis matches. The variety and amount of data made available by these tools are extensive. It surpasses by far the data that can be gathered by manual means. Furthermore, the precision with which this data is collected is considerable and cannot be matched by those means. All of this is made possible by the use of computer vision and a wide range of artificial intelligence models and machine learning algorithms. As these apps are essentially low-cost tools, any non-professional tennis player can have their skills, abilities and capacity evaluated. They do not even need to participate in tournaments.

By classifying non-professional tennis players into categories, these apps can provide a clear vision of the progress that players have made so far. In addition, they can provide suggestions on how they could improve their skills. Furthermore, one can choose to make suggestions selectively with respect to tactical abilities, technical skills, and physical capacity.

Therefore, the model presented in this article, combined with an app to monitor and analyze tennis matches, can be used as an efficient coaching assistant. It helps coaches to devise a series of exercises that their athletes should execute to improve their performance more effectively.

5. Future Work

At the same time as presenting a mathematical model that guides the training activities of tennis players, it also provides a basis for future research opportunities.

For example, an experiment involving a large number of non-professional tennis players can be used to evaluate the model's precision and accuracy. This offers an opportunity to compare the opinion of the experts who helped to develop the model presented in this article with empirical data. Such an experiment would allow the weights attributed to criteria and sub-criteria to be adjusted so that the model's classification accuracy and precision are maximized.

Moreover, a large amount of data may indicate that a similar model must be developed to help very young tennis players improve their performance, as they are still developing motor coordination, strength, and tactical awareness, and might have specific needs and expectations.

Sports-monitoring apps are becoming increasingly available. It is worth investigating how the same steps presented in this article can be used to classify athletes from related sports into categories and support the development of training programs that help to improve their performance.

ACKNOWLEDGEMENTS

We thank the professionals listed below for actively supporting the development of the model presented in this article. Clayvert Gusmão, Isac Gomes, Alexandre Cury and Jorge Werneck Allen played a key role in the construction of the model, while Bruno Cruz played a supporting role.

- **Clayvert Gusmão** is a prominent tennis player. He was number 1 in the Brazilian veterans ranking for 15 consecutive years (2007-2021) and number 1 in the world veteran ranking in 2009. He led the South American ranking for 7 years and is a three-time South American individual champion. He has won 331 titles in 6 countries and has attended renowned tennis education institutions such as the Sanchez-Casal Academy, the Brazilian Tennis Confederation (CBT) and the United States Professional Tennis Registry (USPTR). He has been working as a professional tennis coach for the last 30 years. Gusmão, a graduate in Business Administration, is well known for organizing tournaments and developing the art of tennis in Brazil.
- **Bruno Cruz** is a tennis enthusiast who has been training to become a professional coach with the Brazilian Tennis Confederation (CBT). Currently, he is interning under Clayvert Gusmão's supervision to gain the necessary experience to become a successful tennis coach. He holds a B.Sc. in Actuarial Sciences and has been

advising Gusmão on the use of advanced statistical methods to analyze tennis matches and increase players' performance.

- **Alexandre Cury** is a distinguished tennis player and coach. As a youth player, he was the Rio de Janeiro state champion in the under-12 (US12), under-14 (US14) and under-16 (US16) categories, runner-up in the Brazilian junior singles championship and national junior doubles champion. By the end of his junior career, he was ranked in the top 10 in Brazil in under-18 doubles and in the top 12 in singles. In the United States, he won the Kansas high school state championship representing Iola High School in 1998. From 2001 to 2005, he played for Francis Marion University ranked in the top 6 in NCAA Division 2 at the time. Cury began coaching at age 17 with children in Iola and spent subsequent summers working at Pine Forest Camp (Pennsylvania) and Point O'Woods (Long Island), giving private lessons and conducting clinics for juniors and adults. In Brazil, he has been teaching social tennis for the past six years.
- **Isac Gomes** is a professional tennis coach for high-performance players with almost a decade of professional experience. He started playing tennis at a very early age. Gomes was ranked number 1 in Argentina in the under-16 category (US16) and number 3 in the under-18 category (US18). He played interclub competitions in France, Italy, Germany, and Switzerland for 8 years. He holds international coaching certifications from the Sánchez-Casal Academy and Tennis Europe (the European tennis federation). For 12 years he trained at the Tandil Tenis Club (TTC) in Argentina, where he worked alongside players like Juan Mónaco, Juan Martín del Potro, Diego Junqueira, and Gabriela Sabatini.
- **Jorge Werneck Allen** is an experienced tennis player who has been playing the sport for the past 17 years. He has participated in both the Brazilian circuit and international tournaments in France. A lifelong dedication to tennis has driven his pursuit of excellence in understanding the game and guiding others in their development. Allen holds a bachelor's degree in Civil Engineering, a master's degree in Finance, and is currently pursuing a PhD in Economics at the Getulio Vargas Foundation, a leading educational and research institution in Brazil.

REFERENCES

[1] R. P. Porreca, "General Managers and the Importance of Using Analytics," The Sport Journal, 15 September 2016.

[2] C. Cain, The Sports Statistics Manual: Baseball, Football and Basketball, Amazon, 2024.

[3] A. Phatak, "Logistic Regression," in Computer Science in Sport, Springer, 2024, pp. 135-140.

[4] T. Duan, Z. He, J. Dai, L. Xie, Y. Shi, L. Chen, J. Song, G. Li and W. Zhang, "Effects of unilateral and bilateral contrast training on the lower limb sports ability of college basketball players," Frontiers in Physiology, vol. 15, November 2024.

[5] S. M. Choi, F.-J. Wang, R. K. W. Sum, B. H.-H. Ching, E. F. L. Leung and R. W. K. Ho, "The motivational impact of sport education model on daily physical activity levels among university students: a mediation analysis," Scientific Reports, vol. 14, November 2024.

[6] H. N. Ghani, A. Rusdiana, H. Subarjah, A. M. Syahid and T. Kurniawan, "Dominant Components of Technical Aspects in Identification of Football Athlete Talent Using AHP," Jurnal Ilmu Keolahragaan, vol. 7, no. 2, 2024.

[7] G. S. Nikolovaa, A. B. Dimitrova and D. M. Dantchev, "Mathematical modelling of basic anthropometric and mass inertial characteristics of young tennis players versus nonplayers: I case study for Bulgarian boys," Series on Biomechanics, vol. 37, no. 3, pp. 19-25, 2023.

[8] M. Ma, "Momentum Dynamics in Competitive Sports: A Multi-Model Analysis Using TOPSIS and Logistic Regression," arXiv, p. September, 2024.

[9] A. Khalil, T. Hussain and H. Jamshaid, "Multi-criteria optimization of sports compression socks using Taguchi-VIKOR statistical approach," Journal of Engineered Fibers and Fabrics, vol. 19, pp. 1-15, 2024.

[10] P. An, "Fuzzy Decision Support Systems to Improve the Effectiveness of Training Programs in the Field of Sports Fitness," International Journal of Computational Intelligence Systems, vol. 17, 2024.

[11] S. Sanjaykumar, S. Natarajan, P. Y. Lakshmi and F. A. Boby, "Predicting Team Success in the Indian Premier League Cricket 2024 Season Using Random Forest Analysis," Physical Education Theory and Methodology, vol. 24, no. 2, 2024.

[12] L. Shi, "Research on score prediction in sports competitions ——based on PSO-BP neural network and ARIMAX hybrid prediction model," in Int. Conference on Management Economics and Accounting (ICMEA, Berlin, Germany, 2024.

[13] P. C. Dindorf, E. Bartaguiz, F. Gassmann and M. Fröhlich, Artificial Intelligence in Sports, Movement, and Health, Springer, 2024.

[14] W. Du, "The Computer Vision Simulation of Athlete's Wrong Actions Recognition Model Based on Artificial Intelligence," IEEE Access, vol. 12, pp. 6560-6568, 2024.

[15] G. M. Migliaccio, J. Padulo and L. Russo, "The Impact of Wearable Technologies on Marginal Gains in Sports Performance: An Integrative Overview on Advances in Sports, Exercise, and Health," Applied Sciences, vol. 14, no. 15, 2024.

[16] Slashdot, "Best Sports Video Analysis Software of 2024," Slashdot Media, 2024. [Online]. Available: https://slashdot.org/software/sports-video-analysis/. [Accessed 14th December 2024].

[17] R. Saaty, "The analytic hierarchy process—what it is and how it is used," Mathematical Modelling, vol. 9, no. 3-5, pp. 161-176, 1987.

[18] B. Jones, "Making holistic pairwise judgements," in Academic Standards in Higher Education, Routledge, 2024, pp. 217-226.

[19] J. Ramík, Pairwise Comparisons Method, Springer, 2020.

[20] J. Mazurek, Advances in Pairwise Comparisons, Springer, 2023.

[21] V. S. Prasad and P. Kousalya, "Role of Consistency in Analytic Hierarchy Process – Consistency Improvement Methods," Indian Journal of Science and Technology, vol. 10, no. 29, 2017.

[22] A. Ishizaka and A. Labib, "Review of the main developments in the analytic hierarchy process," Expert Systems with Applications, vol. 38, no. 11, pp. 14336-14345, 2011.

[23] M. Brunelli, Introduction to the Analytic Hierarchy Process, Springer, 2015.

[24] G. Blokdyk, Analytic Hierarchy Process A Complete Guide, STARCooks, 2020.

[25] S. Bozóki, L. Csató and J. Temesi, "An application of incomplete pairwise comparison matrices for ranking top tennis players," European Journal of Operational Research, vol. 248, no. 1, pp. 211-218, 2016.

[26] L. Csató, "A characterization of the Logarithmic Least Squares Method," European Journal of Operational Research, vol. 276, no. 1, pp. 212-216, 2019.

[27] József Temesi, Z. Szádoczki and S. Bozóki, "Incomplete pairwise comparison matrices: Ranking top women tennis players," Journal of the Operational Research Society, vol. 75, no. 1, pp. 145-157, 2024.

[28] G. G. P. Garcia and J. C. M. Mori, "Who's the Goat? Sports Rankings and Data-Driven Random Walks on the Symmetric Group," arXiv preprint, p. arXiv:2409.12107, 2024.

[29] W. Wu, "Quantification of momentum in tennis matches and its impact: a study based on AHP-EWM method and data analysis," Highlights in Science, Engineering and Technology, vol. 100, pp. 142-149, 2024.

[30] J. Frost, Hypothesis Testing: An Intuitive Guide for Making Data Driven Decisions, Amazon, 2020.

[31] M. Harrison, Effective XGBoost: Optimizing, Tuning, Understanding, and Deploying Classification Models, MetaSnake, 2023.

[32] G. C. Uribe, Genetic Algorithms Made Simple: A Beginner's Guide to Solving Optimization Problems, Amazon, 2024.

[33] S. Kang, Y. Shi, Y. Chen and H. Wang, "Enhancing Tennis Match Predictions with AHP, XGBoost, and Genetic Algorithms," in 2024 IEEE 4th International Conference on Software Engineering and Artificial Intelligence (SEAI), Xiamen, China, 2024.

[34] H. Mladen, N. Đurović and L. Stanišić, "Priorities and Rankings for Offensive Baseliner Tennis Player," in 9th International Scientific Conference on Kinesiology, Opatija, Croatia, 2021.

[35] J. Zhang, Q. Zhu, L. Xu and H. Gong, "Research on tennis match evaluation and prediction model based on data mining," in Fourth International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC), Kaifeng, China, 2024.

[36] R. Kumar, S. Singh, P. S. Bilga, J. J. Singh, Sunpreet, M.-L. S. Scutaru and C. J. Pruncu, "Revealing the benefits of entropy weights method for multi-objective optimization in machining operations: A critical review," Journal of Materials Research and Technology, vol. 10, no. January–February, pp. 1471-1492, 2021.

[37] P. Bloem, Unraveling principal component analysis, Amazon, 2023.

[38] H. v. Maarseveen, Support Vector Machines: The Groundbreaking Model That Changed Machine Learning, Amazon, 2023.

[39] Y. Lei, A. Lin and J. C. Jr., "Rhythms of Victory: Predicting Professional Tennis Matches Using Machine Learning," IEEE Access, vol. 12, pp. 113608-113617, 2024.

[40] ESPN, "Men's Tennis ATP Rankings 2025," 2025. [Online]. Available: https://www.espn.co.uk/tennis/rankings. [Accessed 23 June 2025].

[41] Z. Guangfu and T. Kanchanathaweekul, "Selected Indicators Influence the Talent Identification of Young Tennis," International Journal of Sociologies and Anthropologies Science Reviews (IJSASR), no. November-December, pp. 229-246, 2023.

[42] ImagineLabs, "Think science. Drive innovation," [Online]. Available: https://imaginelabs.com.br/. [Accessed 10 01 2025].

[43] totaltenis.ai, "Total Tennis," [Online]. Available: https://litcomp.cbpf.br/dev/totaltennis/login. [Accessed 10 01 2025].

[44] F. D. Felice and A. Petrillo, "Absolute measurement with analytic hierarchy process: A case study for Italian racecourse," International Journal of Applied Decision Sciences, vol. 6, no. 3, pp. 209 - 227, 2013.

[45] T. L. Saaty, "Absolute and Relative Measurement with the AHP. The most Livable Cities in the United States," Socio-Economic Planning Sciences, vol. 20, no. 6, pp. 327 - 331, 1986.

[46] F. D. Felice, V. Elia, M. Gnonib and A. Petrillo, "Comparing environmental product footprint for electronic and electric equipment:a multi-criteria approach," International Journal of Sustainable Engineering, vol. 7, no. 4, p. 360–373, 2014.

[47] T. R. Ryden, Occam's Razor, Beaufort Books, 2019.

[48] G. C. Uribe, Genetic Algorithms Made Simple: A Beginner's Guide to Solving Optimization Problems, Amazon, 2024.

[49] J. H. Holland, Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, MIT Press, 1975.