RESEARCH ARTICLE

Evaluate performance trends of professional tennis players using historical ATP rankings and match results

Jiajia Li[1], Xiao Wang[2], Xuehao Xia[3, *]

[1]Woosuk University, Jeonju, South Korea. [2]Zhejiang Changzheng Vocational & Technical College, Hangzhou, Zhejiang, China. [3]Physical Education College, Huangshan University, Huangshan, Anhui, China.

Received: March 10, 2025; **accepted:** June 23, 2025.

Professional tennis is a rapidly growing sport. Analyzing trends in individual player careers can lead to effective strategies for choosing and planning the team. Association of Tennis Professionals (ATP) data includes scores and rankings, allowing coaches to understand player performance and predict their future performance. This research assessed the performance of tennis professionals throughout the years by analyzing their ATP ratings and match data to identify a tennis player's capacity for consistency, peak performance during certain games, adaptability to a variety of tactics, ability to respond to injuries or tournament victories using a framework based on machine learning. Publicly accessible ATP records were used to create dataset, which included player ratings, match outcomes, and tournament finishes. During data preprocessing, the normalization, fixing missing data, and transforming were important steps for consistency. The recursive feature elimination (RFE) was used to choose important features, looking at trends in win-loss records, ranking changes, and the winner in each matchup. The approach introduced a hybrid model of weighted coyote optimization algorithm (WCOA) and gradient boosting decision tree (GBDT) as WCOA helping to improve the performance of GBDT. By using the proposed WCOA-GBDT model, it was possible to achieve a prediction accuracy of 97.10%, together with precision of 95.20%, recall of 94.60%, F1-score of 93.40%, R^2 of 0.991, mean absolute percentage error (MAPE) of 1.110 and root mean square error (RMSE) of 0.033. The results supported that the model performed well at forecasting future ATP ratings and identifying promising talent. This analysis showed the specific and data-driven features of professional tennis performance. The results showed that doing well in the past and having the ability to change were the keys to maintaining exceptional performance. This research was important for coaches, analysts, and stakeholders who planned to develop top athletes and guide elite players.

Keywords: machine learning; Association of Tennis Professionals; recursive feature elimination; weighted coyote optimization algorithm; gradient boosting decision tree; playing styles.

[*]**Corresponding author:** Xuehao Xia, Physical Education College, Huangshan University, Huangshan, Anhui 245041, China. Email: 15105590588@163.com.

Introduction

Sports activities can enhance both physical and mental benefits. Participating in sports regularly strengthens the heart, builds muscles, increases flexibility, improves coordination, and builds endurance [1, 2]. Apart from helping human bodies, sports also influence human minds by making people feel less stressed, lifting spirits with hormones, and boosting attention and balance of feelings [3]. Playing sports can teach people to manage time, develop teamwork, increase self-discipline, strengthen mental fitness, and help them build self-esteem and social connections [4]. Analyzing performance patterns is now considered important to help understand the things that influence players' lasting achievements in tennis [5, 6].

By using Association of Tennis Professionals (ATP) rankings and tournament results, researchers can learn how athletes improve, perform at their peak, and how long their careers persist. These indicators help evaluate physical readiness, mental power, and the ability to adjust and are all important for enduring success [7]. In addition, statistics about player and team performance are now very valuable to the sports betting industry since they guide important choices, better training systems, and the evaluation of competitors [8]. Players are ranked according to ATP based on their results in matches, number of victories, game results, and tournament position [9]. Mainer-Pardos *et al*. proposed a system to compare one player's performance with others and tracked their consistency using match statistics and rankings, which made it easier to find trends and plan for future outcomes. However, there was still a gap in how consistently, strategically, and resiliently players deal with various tournament environments [10]. Oliver *et al*. examined the pertinent factors that contributed to professional tennis players' retirement from match play and found that retirement rates had been increasing due to match-related reasons such as the type of court used, the tournament level, and how far the player advanced in the draw [11]. Fayomi *et al*. proposed the Bradley-Terry model and analyzed 3,439 matches played in the year and a half after January 2019. The results showed that the model often performed best on clay courts, surpassing the ATP rankings and bringing investors better profit margins than bookmakers' odds [12]. Rather than relying just on the outcome of each match, Almarashi *et al*. proposed a model to forecast tennis outcomes through both historical matches and their scheduling using both linear and nonlinear time series models and achieved strong performance in projecting player performance [13]. Further, Breznik *et al*. analyzed 30 years data of professional tennis and found that strong and non-strong pairs usually played for a short amount of time, while strong players were often engaged in longer battles [14]. Perri *et al*. investigated the tournament schedules of prospective top-100 and top-250 male tennis players between the ages of 13 and 18 for 165 players using volumes, inter-tournament intervals, and consecutive events and found that teenagers' participation in tournaments and matches increased sharply at age 15 and future top-100 players entered professional events sooner with this group tending to play more matches consistently [15]. Another research suggested that, by using cluster optimization to design a balanced drawing of candidates, it could improve the schedule fairness compared to traditional solutions in single-elimination tennis tournaments [16]. Arcagni *et al*. proposed a logistic regression model using eigenvector centrality to update player ratings after every match and found that this model did a better job of predicting outcomes and offered successful outcomes in wagering [17]. D'Urso *et al*. developed a model to categorize tennis players and tournaments by their similarities and relationship using ATP and Wheelo rating data. The model performed better than the bipartite network approach [18].

Previous research in tennis provides meaningful findings about performance improvement and tournament equity augmentation, together with predictive model enhancements. However, data accessibility and processing still face restrictions. Furthermore, those models do not use historical trends during forecasting. To overcome these challenges, this study aimed to predict and examine professional tennis players' performance by evaluating player performance trends in professional tennis through assessing the machine learning (ML) model of the weighted coyote optimization algorithm gradient boosting decision tree (WCOA-GBDT) using ATP historical data to reflect professional tennis performance dynamics and offer insightful information for forecasting future achievement. This research also identified the important indicators such as a team's win-loss ratio and changes in their ranking using applied recursive feature elimination (RFE). The results of this research would provide actionable insights and practical suggestions for coaches, analysts, and stakeholders in professional tennis to identify talent and design strategies.

Materials and methods

Data resource and collection

The data were obtained from Kaggle database (https://www.kaggle.com/datasets/warcoder/atp-tennis-rankings-results-and-stats1968-2023) with the information of the world ranking of ATP players, individual match results, final tournament outcomes, and statistics collected over five decades. A total of 150,000 match records from 1968 to 2023 was retrieved from the database, which exclusively covered male professional tennis players as it was based on the ATP tour. The data included various types of match records such as Grand Slam matches, ATP Tour events (Masters 1000, 500, 250), Davis Cup matches, and Challenger events along with corresponding ranking data, player statistics, and tournament outcomes.

Data pre-processing

The original data were pre-processed through missing value handling and normalization to ensure consistency across the dataset. Missing value handling made the researchers aware and replace missing information by using K-Nearest Neighbors (KNN) algorithm to impute player statistics and ranking points that were missing in the data. When dealing with variable types like tournament results and injury statuses, the most common value was used to fill in missing data. To replace missing ATP position rankings, linear interpolation was used to ensure that the data remained continuous in time. After data pre-processing, there were no missing values in the data used to train the WCOA-GBDT model. The data were then normalized using min-mix normalization to maintain the connections among the initial information of tennis data, which was an easy method to correct the position inside a predefined boundary as follows.

$$B' = \left(\frac{B - \min value\ of\ B}{\max value\ of\ B - \min value\ of\ B}\right) * (C - D) + D \qquad (1)$$

where *B′* was one among the min-max standardized sets of information being used to evaluate player performance trends in professional tennis. *B* was subsequently converted tennis data. *C* and *D* were the predefined perimeters if *B* was the starting region.

Feature extraction

Recursive feature elimination (RFE) functioned as a feature selection technique that enhanced model performance through repeated elimination of the least important features during its iteration method was employed to evaluate player performance trends in professional tennis. The least important feature was eliminated for each iteration, which ran until an optimal set of retained features was determined. RFE decreased dimensionality and boosted model performance alongside boosting interpretability as follows.

$$z = e\ (A, b) \qquad (2)$$

where *z* determined the features coefficient. *A* was the matrix features. *b* was the variables of the target. *e* was utilized to assess the tennis data.

Assessing ATP matches by WCOA-GBDT model

The model WCOA with GBDT evaluated the ATP tennis match results and, as the predictive model, integrated optimization methods with ML to optimize forecasting accuracy. This combination improved both the precision and operational speed in forecasting match results. The gradient boosting decision tree (GBDT) constituted an effective and visually appealing method to provide an organized depiction of potential results and the choice routes leading in tennis data. The twigs kept growing and creating something such as a living thing that assisted in clearly and methodically illustrating the process of making choices in tennis data. The membership values were taken as a sample as shown below.

$$E(H\left(d_{n_j}(w)\right) = \sum_{j=1}^{i=o} d_{n_l}^{i}(w) \quad (3)$$

This calculation valued the energy or aggregation for the data node d_{n_j} in the decision tree as the total effect of samples reaching that node. Consequently, it was possible to determine the energy of the data level nodes on levels below.

$$FH\ (H) = \sum_{n=j}^{n=l} \frac{E\left(H\left(d_{n_j}(w)\right)\right)}{\sum_{n=1}^{n=l}(EH(d_n(w)))} log_2 \frac{E\left(H\left(d_{n_j}(w)\right)\right)}{\sum_{n=1}^{n=l}(EH(d_n(w)))} \quad (4)$$

The entropy or information measure $FH\ (H)$ for the data at a specific tree level was calculated by adding log probabilities of normalized energy values through multiplication. By calculating how impure or uncertain the nodes at that level, it helped decide the splits to use. The flexible segmentation of feature trees and the characteristic node-specific computation were shown below.

$$FH\left(\frac{H_j}{B}\right) = -\sum_{n=1}^{n=l} \frac{E\left(B_j(d_n(w)\right)\cap E(d_n(w)))}{\sum_{n=1}^{n=l} E(B_j(d_n(w))} \ FH\ (H \cap B_j) \quad (5)$$

Equation (5) showed how much information was lost in the feature subset H_j, given the partition B_j. It checked the current uncertainty in the data H_j after selecting the subset B_j to rate the goodness of a particular node split. The relevant data was then obtained as follows.

$$Again\ (B_j, H) = FH(H) - FH(\frac{H}{B_j}) \quad (6)$$

The equation (6) calculated the benefits of splitting data based on B_j's approach. Splitting the dataset decreased entropy, which helped choose the optimal feature and threshold for the node in the tree. The weighted coyote optimization algorithm (WCOA) enhanced the performance of ML models by optimizing feature selection and model parameters to evaluate player performance trends in professional tennis. It improved prediction accuracy by efficiently searching for optimal solutions in tennis datasets as follows.

$$SAC_d^{h,s} = z = [z_1, z_2, \dots, z_C] \quad (7)$$

where *SAC* was the social adaptation condition of the d^{th} coyote from the h^{th} pack during the s^{th} iteration. For every variable z, there was a dimension that handled and represented either feature data or items from the model.

$$SAC_{d,i}^{h,s} = KA_i + l \times (Gq_i - Kq_i) \quad (8)$$

Equation (8) established a new solution for the candidate using knowledge acquisition KA_i, random learning factor l, and the difference between two knowledge quantifiers Gq_i and Kq_i. It mimicked the kind of learning that came from social activities.

$$obj_d^{h,s} = e\left(SAC_{d,i}^{h,s}\right) \quad (9)$$

Equation (9) estimated the value of $e\left(SAC_{d,i}^{h,s}\right)$ for the learned *SAC* to see if it was a prediction error and checked if the solution was fit. At the WCOA graduation, every WCOA participated at random from sets. Additionally, each person modified the condition by switching to the other sets and was expressed as below.

$$O_k = \frac{5}{100} \times M_d^2 \tag{10}$$

The equation regulated how much the number of coyotes M_d^2 changes through offspring or mutation based on the need for diversity. Alpha coyote was the coyote leader of all the sets and was renowned for having high levels of responsibility as shown in equations (11) and (12).

$$\alpha_d^{h,s} = sac_d^{h,s} forminob^{h,s} i_d \tag{11}$$

$$cul_i^{h,s} = \begin{cases} Q_{\frac{M_D+1}{2}}^{h,s}, & M_d \text{ is an odd number} \\ \frac{1}{2}\left(Q_{\frac{M_D}{2},i}^{h,s} + Q_{\frac{M_D+1}{2},i}^{h,s}\right) & P.X \end{cases} \tag{12}$$

With the help of median or average pack positions, it determined how the cultural tendency $cul_i^{h,s}$ was computed, depending on whether M_d was even or odd. The WCOA also considered a collection of environmental elements, the social attitude, and the life process of coyotes as shown in equations (13) to (15).

$$Ble_i^{h,s} = \begin{cases} sac_{l_1,i}^{h,s}, & l_i < hl_t orj = i_1 \\ sac_{l_2,i}^{h,s}, & l_i \geq hl_t + hl_b orj = i_2 \\ \rho i, & P.X \end{cases} \tag{13}$$

The equations illustrated the genetic logic $Ble_i^{h,s}$ with thresholds hl_t and hl_b that governed mating selection and a randomness factor ρi.

$$hl_t = \frac{1}{c} \tag{14}$$

$$hl_b = \frac{1}{2}(1 - hl_t) \tag{15}$$

The set thresholds were established to check if a steer was allowed to breed using the constant c. The replacement of culture among sets was determined using μ_1 *and* μ_2 as shown in the equations (16) to (20).

$$\mu_1 = \alpha_d^{h,s} - sac_{d1}^{h_j s} \tag{16}$$

$$\mu_2 = cul^{h,s} - sac_{d2}^{h_j s} \tag{17}$$

The SAC with changes from the steer's l_1 and l_2 deviation vectors was updated as follows.

$$msac_d^{h,s} = sac_d^{h,s} + l_1 \times \mu_1 + l_2 \times \mu_2 \tag{18}$$

$$mobj_d^{h,s} = e\left(msac_d^{h,s}\right) \tag{19}$$

$$sac_d^{h,s+1} = \begin{cases} msac_d^{h,s}, & mobj_d^{h,s} < obj_d^{h,s} \\ sac_d^{h,s}, & P.X \end{cases} \tag{20}$$

where $mobj_d^{h,s}$ was the goal for the new *SAC* vector. The new role ($msac_d^{h,s}$) was chosen to take only if it made the solution better. The ability of this approach to avoid becoming trapped in a local ideal was one of its primary characteristics. To get a higher classification efficacy, the WCOA approach generated an ff and defined a positive integer to demonstrate the best outcomes of the potential solution. At this stage, ff was assumed that the classifier error rate declined as shown below.

$$fitness(w_j) = Classifier\ Error\ Rate\ (w_j) = \frac{No.Of\ misclassified\ samples}{Total\ No.of\ samples} * 100 \tag{21}$$

where the classifier measured the error on feature subset $fitness(w_j)$, which WCOA tried to reduce as much as possible. The WCOA-GBDT model used gradient boosting together with WCOA to select optimal features while enhancing its prediction of rankings. A weight adjustment system allowed this approach to optimize key statistical and match outcome variables for enhanced performance evaluation. Through this combined framework, better decision outcomes in player development and team strategy functions could be achieved.

Validation of the model

The proposed model was designed and tested on a Python system running on an Intel Core i9 CPU, 8 GB RAM, and Windows 11 for optimal performance. The WCOA-GBDT model was validated in its accuracy, precision, recall, and F1-score using coefficient of determination (R^2), mean absolute percentage error (MAPE), and root mean square error (RMSE). Further, the performance of the model was compared with a Fully Connected Neural Network (FCNN) [19], One-Dimensional Convolutional Neural Network (1D-CNN) [19], and the Cat Boost Regression and Random Forest Algorithm (CBRF) approach [20].

Results and discussion

Comparison of the performance of different models

Accurate ranking predictions were important both for analyzing player performance trends and for spotting emerging talent, which was responsible for guessing how each player would perform in a specific tournament. The accuracy of the model depended on choosing the main factors that could be accuracy, as well as ranking changes and head-to-head statistics. The results showed that the proposed WCOA-GBDT model achieved a higher accuracy of 97.10% than FCNN's 92.6% and 1D-CNN's 93.35% (Figure 1A). Precision focused on main factors like the ratio of wins to losses, changes in ratings, the outcome of mutual matches and reviews about taking part in tournaments and recovering from injuries. To achieve high success, a player must combine excellent previous results and good conditioning with the ability to get used to different playing patterns. The proposed WCOA-GBDT showed the highest precision of 95.20% compared to FCNN of 92.19% and 1D-CNN of 92.30% (Figure 1B). Recall was very useful in evaluating how well the model did at recognizing the most talented players and anticipating their rankings. In this research, ATP rankings were evaluated by reviewing if new players were spotted correctly, even if they changed their positions on the rankings. Despite reversals, insights about professional tennis could still emerge as the use of recall method to spot talented young players who might present stakeholders with useful opportunities for predicting future results in professional tennis. The recall achieved by WCOA-GBDT was 94.60%, which was higher than the recalls of FCNN as 91.54% and 1D-CNN as 91.63% (Figure 1C). The F1 score established itself as vital for assessing professional tennis player performance patterns by producing an optimal judgment between match outcome prediction accuracy and correct ranking predictions. The proposed WCOA-GBDT predictive model employed the F1-score as the metric to evaluate rising talent identification and future ranking prediction. The combination of precision and recall performance appeared in the F1-score, which determined how accurately the model identified player movements without producing many unsatisfactory predictions in the same sample set. The results showed that the F1-scores of WCOA-GBDT, FCNN, 1D-CNN were 93.40%, 91.79%, 91.82%, respectively (Figure 1D).

Comparison of the performance metrics of models

The evaluation of player performance trends in ATP rankings and match outcomes allowed investigators to assess past data for understanding success elements in tennis by R^2. The R^2 revealed important findings about stability

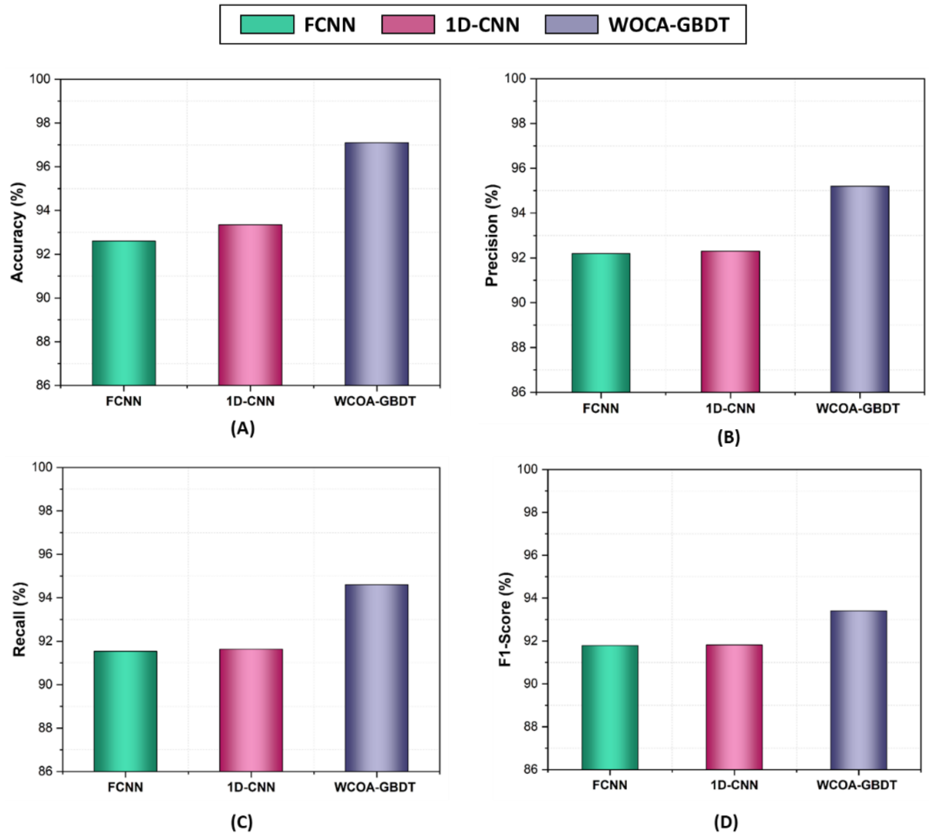

Figure 1. Comparison of the performance for assessing player performance trends in ATP tennis. **A.** accuracy. **B.** Precision. **C.** Recall. **D.** F1 score.

and response abilities and developmental patterns through the evaluation of ranking data alongside player's statistics and enabled the identification of current and prospective athlete development alongside future ranking positions. Therefore, professionals in the tennis data industry could make informed strategic decisions. The R^2 of WCOA-GBDT was 0.991, while that of the existing CBRF method was 0.985 (Figure 2A). The MAPE served as an essential element for performance trend evaluation by measuring how accurately the predictive models represented their intended predictions, which helped evaluate predictions against actual ranking outcomes in ATP events. MAPE enabled performance evaluation of predictive models by determining percentage differences between actual and forecast results, thus enabling better precision for validated predictions. The results showed that the accurate player trajectory modeling was 1.110 for WCOA-GBDT and 2.185 for CBRF (Figure 2B). The RMSE was essential for assessing tennis player performance because it regulated the accuracy of ATP rankings prediction models. The research used RMSE as an assessment tool to find quantitative gaps between forecasting and observed player rankings to evaluate model prediction accuracy. The hybrid WCOA-GBDT model demonstrated its capability to track player outcomes through RMSE values that indicated precise forecasting ability with WCOA-GBDT at 0.033 and CBRF at 0.055 (Figure 2C).

The model FCNNs showed a tendency to over-fit during processing small or imbalanced data samples along with encountering high computational requirements, while 1D-CNNs came with limitations to analyze sequential data in the recognition of tennis performance complexities and did not provide clear insights into individual feature importance [19]. CBRF

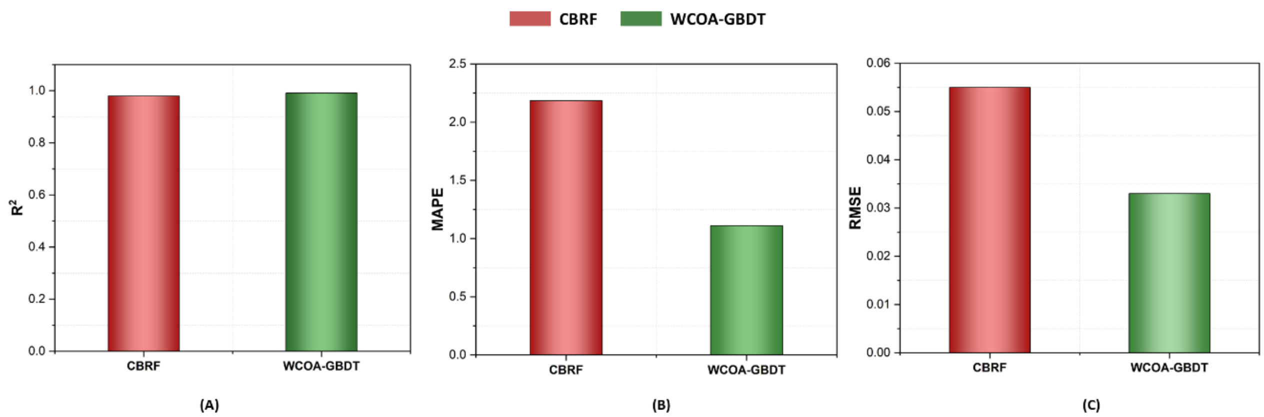

Figure 2. Comparison of performance metrics of models for tennis player trends evaluation.

model faced issues with understanding nonlinear patterns between variables and frequently adopted characteristics that should be ignored during training [20]. The proposed WCOA-GBDT algorithm reduced over-fitting features by ensuring better generalization of models. Through its mechanism, the system provided complete clarity about the elements that affected player rankings and performance patterns. The WCOA-GBDT of the ML model showed strong capability for future ranking prediction as it considered past performance along with consistency and adaptability as essential factors for enduring achievement. The results confirmed that the proposed WCOA-GBDT had effectively evaluated the player performance trends in professional tennis and achieved accuracy of 97.10%, precision of 95.20%, recall of 94.60%, F1-score of 93.40%, R^2 of 0.991, MAPE of 1.110, and RMSE of 0.033. However, some crucial elements such as the player's psychology and specific training regimens were not examined because they were only found in publicly available records. Future research developments should include extensive performance data acquisitions alongside real-time statistics and medical records of injuries together with additional factors that affect player development to achieve enhanced model accuracy.

References

1. Zappalà C, Biondo AE, Pluchino A, Rapisarda A. 2023. The paradox of talent: How chance affects success in tennis tournaments. Chaos Soliton Fract. 176:114088.
2. Bozděch M, Zháněl J. 2023. Analyzing game statistics and career trajectories of female elite junior tennis players: A machine learning approach. PLoS One. 18(11):e0295075.
3. Wu M, Fan M, Hu Y, Wang R, Wang Y, Li Y, *et al*. 2022. A real-time tennis level evaluation and stroke classification system based on the Internet of Things. Internet Things. 17:100494.
4. Néri-Fuchs JB, Sedeaud A, Marc A, De Larochelambert Q, Toussaint JF, Brocherie F. 2023. Medical withdrawals in elite tennis in reference to playing standards, court surfaces, and genders. J Sci Med Sport. 26(6):296-300.
5. Pineda-Hernández S. 2022. Playing under pressure: EEG monitoring of activation in professional tennis players. Physiol Behav. 247:113723.
6. Wang H, Yang J, Hu M, Tang J, Yu W. 2023. A comparative analysis for eye movement characteristics between professional and non-professional players in FIFA eSports game. Displays. 81:102599.
7. AlShami A, Boult T, Kalita J. 2023. Pose2Trajectory: Using transformers on body pose to predict tennis player's trajectory. J Vis Commun Image Represent. 97:103954.
8. Pineda-Hernández S. 2022. How to play under pressure: EEG monitoring of mental activation training in a professional tennis player. Physiol Behav. 250:113784.
9. Zhong M, Liu Z, Liu P, Zhai M. 2024. Searching for the effects of momentum in tennis and its applications. Procedia Comput Sci. 242:192-199.
10. Mainer-Pardos E, Álvarez VEV, Moreno-Apellaniz N, Gutiérrez-Logroño A, Calero-Morales S. 2024. Effects of a neuromuscular training program on the performance and inter-limb asymmetries in highly trained junior male tennis players. Heliyon. 10(5):e27081.
11. Oliver L, Baiget E, Cortés J, Martínez J, Crespo M, Casals M. 2024. Retirements of professional tennis players in ATP and WTA tour events. Eur J Sport Sci. 24(10):1526-1536.

12. Fayomi A, Majeed R, Algarni A, Akhtar S, Jamal F, Nasir JA. 2022. Forecasting tennis match results using the Bradley-Terry model. Int J Photoenergy. 2022(1):1898132.
13. Almarashi AM, Daniyal M, Jamal F. 2024. A novel comparative study of NNAR approach with linear stochastic time series models in predicting tennis player's performance. BMC Sports Sci Med Rehabil. 16(1):28.
14. Breznik K, Candila V, Milekhina A, Restaino M. 2024. Eras of dominance: Identifying strong and weak periods in professional tennis. Comput Stat. 2024:1-18.
15. Perri T, Duffield R, Murphy A, Mabon T, Reid M. 2021. Competition scheduling patterns of emerging elite players in professional men’s tennis. J Sports Sci. 39(18):2087-2094.
16. Della Croce F, Dragotto G, Scatamacchia R. 2022. On fairness and diversification in WTA and ATP tennis tournaments generation. Ann Oper Res. 2022:1-13.
17. Arcagni A, Candila V, Grassi R. 2023. A new model for predicting the winner in tennis based on the eigenvector centrality. Ann Oper Res. 325(1):615-632.
18. D’Urso P, De Giovanni L, Federico L, Vitale V. 2024. Network and attribute-based clustering of tennis players and tournaments. Comput Stat. 2024:1-24.
19. Domínguez GC, Álvarez EF, Córdoba AT, Reina DG. 2023. A comparative study of machine learning and deep learning algorithms for padel tennis shot classification. Soft Comput. 27(17):12367-12385.
20. Lv X, Gu D, Liu X, Dong J, Li Y. 2024. Momentum prediction models of tennis match based on CatBoost regression and random forest algorithms. Sci Rep. 14(1):18834.