COACHING & SPORT SCIENCE REVIEW

International Tennis Federation
www.itfcoachingreview.com
April 2024. 32th Year. Issue 92. 46-48
ISSN 2225-4757
https://doi.org/10.52383/itfcoaching.v33i92.563

# What can artificial intelligence do for tennis?

Fernando Vives Albiol

Valencia University, Spain.

**ABSTRACT**

In the current era of Artificial Intelligence, we are witnessing how this technology is revolutionizing the world of sports. Through a review of the main Machine Learning research in tennis over the last decade, players, coaches, and fitness trainers can discover new proposals to improve and personalize training sessions, enhance player effectiveness, and optimize decision-making during competition.

**Key words:** machine learning, performance analysis, artificial intelligence, research.

**Received**: 12 January 2024

**Accepted**: 19 March 2024

**Corresponding author:** Fernando Vives Albiol. Email: fervial@alumni.uv.es

## INTRODUCTION

In our days, it is increasingly common to talk about the terms Artificial Intelligence (AI) or Machine Learning (ML) in the sports context, but what do they mean? How do they work? What studies have been carried out? What applications can they have in the world of tennis? The term AI was first used by John McCarthy in 1956, who later proposed the following definition: "It is the science and engineering of creating intelligent machines, especially intelligent computer programs. It is related to the task of using computers to understand human intelligence, but AI does not have to be limited to biologically observable methods." In other words, AI is characterized by the combination of computer science and data analysis to address complex problems (McCarthy, 2004). Russell and Norvig (2010) contributed with a renewed approach to the study of AI, classifying computer systems according to their ability to reason and act. ML, is a subset of AI conceptualized initially by Arthur Samuel in 1959, which allows computers to improve in specific tasks without using explicit programming. Essentially, ML uses computer algorithms to analyze data and learn from it through experience, classifying or predicting a certain event (Mitchell, 1997). The development of an ML model consists of the following parts:

- Selection and preparation of the dataset
- Choice of algorithm or set of algorithms
- Model training
- Use and improvement of the model

Deep Learning (DL) constitutes a subset within the field of Machine Learning (ML). Among all DL algorithms, neural networks stand out above all, which emulate the biological neurons of the human being, simplifying their operation and focusing on information processing. These Artificial Neural Networks (ANNs) have shown great effectiveness in solving classification, prediction, optimization, or pattern recognition problems (Stanko, 2020; Thakur & Konde, 2021).

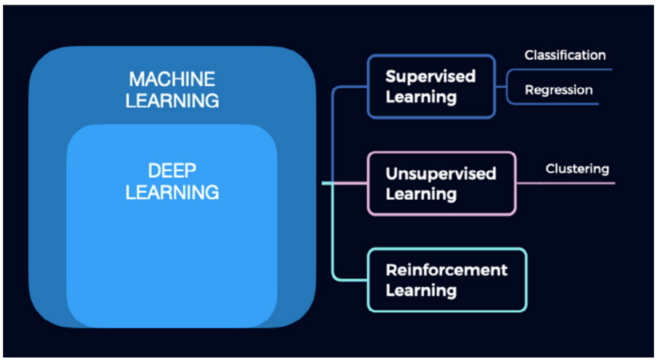


**Figure 1.**

## ML IN TENNIS

In the last two decades, technology has experienced an unstoppable advance, manifesting itself both in everyday life and in the sports field. The collection of data and its quality, through systems such as Tracking or Tagging, has caused a significant change in the landscape of professional sports and in Sports Science research. The emergence of 'big data' in professional tennis, largely facilitated by the incorporation of Hawk-Eye (HE) in 2006, has allowed for more sophisticated analysis through the application of ML techniques and opened up new research approaches associated with tennis (Chase, 2020). In recent years, there has been a significant increase in the number of scientific articles that employ ML techniques in the field of tennis. This increase reflects the recognition of ML's effectiveness in addressing specific challenges in the tennis context, providing new perspectives and innovative approaches to analyzing data, improving players' performance, and better understanding game patterns. Below, we will briefly review different studies that have contributed to this in recent years.

The serve has become a fundamental shot in modern tennis, as evidenced by match statistics in the men's category. Players initiate the point with it and seek to take control of the game. Wei et al. (2015) analyzed 7,050 serves using Hawk-Eye (HE) data during three editions of the Australian Open (2012-2014) to predict the most likely serve of a player in a given context. The aim of the study was to provide coaches and players with a useful tool for match preparation against a specific opponent, which could even be used during the match in a given situation.

The following study also focused on the serve. Data from HE on 262,596 points from Grand Slam (GS), ATP, and WTA tournaments between 2003 and 2008 were used. A comprehensive analysis of first and second serves was conducted, evaluating their importance based on surface, speed, direction, and spin applied (Mecheri et al., 2016). It was one of the first studies to handle a huge amount of data and delve into the serve, highlighting the importance of taking a comprehensive approach to it by providing clear and concise performance indicators in both men's and women's individual tennis. This can provide coaches with indispensable material to prepare their training sessions most efficiently.

In 2017, Whiteside and Reid conducted a study on the most determining characteristics a serve should have to achieve an ace. For this, they analyzed 25,680 first serves from 151 matches in the men's draw of the Australian Open played between 2012 and 2015. The serve angle and bounce distance to the line were decisive in achieving a direct ace, and this information can be crucial for players both to plan training sessions more selectively and to improve decision-making in matches at critical moments.

Another study, this time by Kovalchik and Reid (2018), analyzed the serves, rallies, and points of the Australian Open between 2015 and 2017. The data were collected through the Hawk-Eye system, with a total of 448,159 shots in over 400 matches between men and women. They identified 13 different types of serves in men, while in women, 17 types were identified on the Advantage side and 15 on the Deuce side. They also provided a comprehensive taxonomy of the different types of tennis shots in the male and female individual categories. This information can provide coaches with a very powerful tool to prepare their players more specifically and representatively.

Reading the ball to try to anticipate the opponent is crucial during exchanges in matches. For this reason, Shimizu et al. (2019) proposed a novel method to predict the direction of a player's next shot based on their posture and position prior to hitting the ball. Players could conduct video analysis sessions to study their opponent to predict their shots in a given context in the preparation of their matches, as is often done in other sports.

Stan Wawrinka, after defeating David Ferrer in the 2014 Monte Carlo semifinals, stated: "I know that when I move well, I can dictate the rhythm of the game." The following study by Giles et al. (2020) identified and classified medium and high-intensity change of direction (COD) movements in professional male and female tennis. The speed, distance covered, change of inclination, and acceleration of both male and female players were examined to identify the significant physical demands of such a dynamic sport as tennis. This data is crucial for physical trainers and athletes, as it provides valuable information to improve physical preparation during training sessions, which can translate into more optimal performance in competition.

The volley is another shot that has been studied in depth in tennis research. Thus, Martinez-Gallego et al. (2021) studied the different types of volleys that occur in men's and women's doubles matches belonging to the Davis Cup and the Billie Jean King Cup, respectively. The results showed 7 different types of volleys in the men's category, while only 4 different types were obtained in the women's category. These findings can be very useful, as the volley is a very specific tennis shot, knowing the different types of volleys that occur during competition can make a difference during the game.

The return has also been the subject of study using ML methodology. Kovalchik and Albert (2022) analyzed 142,803 points belonging to 141 male professional players between the years 2018 and 2020. The results showed six different return zones for first serves and six different zones for second serves. Similar to the previous study on volleys, this information can be crucial for planning and specific preparation based on the different types of returns.

As previously mentioned, the serve is a decisive shot in modern tennis in both singles and doubles disciplines. Therefore, Vives et al. (2023) analyzed a total of 14,146 first serves from Davis Cup ties played between 2010 and 2019. The angle of the serve and the distance of the bounce to the sideline were key factors in achieving a direct ace, far more than speed. Hence, players can have very specific parameters to increase their effectiveness on first serves, thus optimizing decision-making for servers in crucial moments of the match.

Lastly, Zhou and Liu (2024) examined the preference for the type of stance in male professional players. The methodology included data analysis from the Australian Open using Bayesian network models, highlighting the predominance of open and semi-open stances in forehand strokes, and the closed stance in two-handed backhand strokes. The results obtained showed that the player's position and the ball's bounce zone determined the player's choice of stance type. Therefore, coaches could undertake much more defined work in hitting zones during training sessions.

## PRACTICAL APPLICATIONS

As we have observed in the previous section, ML has been developing in the field of modern tennis research. The results of the different studies provide very concise and detailed information in different areas of the game, ranging from specific shots such as the serve, return, or volley, to predicting the direction of the next shot, the type of stance, or the physical demands during competition. This information can be very interesting for optimizing and personalizing training programs by coaches and trainers, maximizing the effectiveness of players, and improving decision-making during matches. Given that high-performance matches are increasingly closely contested, often it is the small details that can make the difference between a victory and a defeat.

**Table 1.** *Summary of the main ML studies in tennis.*

| AUTHOR(S) | YEAR | SAMPLE | DATA | AREA OF STUDY |
|---|---|---|---|---|
| WEI ET AL. | 2015 | 4.758 1° SERVES, 2.292 2° SERVES | HAWK EYE | SERVE |
| MECHERI ET AL. | 2016 | 262.596 SERVES | HAWK EYE | SERVE |
| WHITESIDE & REID | 2017 | 25.680 1° SERVES | HAWK EYE | DIRECT SERVE |
| KOVALCHIK & REID | 2018 | 270.023 SHOTS MEN<br>178.136 SHOTS WOMEN | HAWK EYE | SHOT TAXONOMY |
| SHIMIZU ET AL. | 2019 | 1 VIDEO, 1 MATCH | YOUTUBE | SHOT PREDICTION |
| GILES ET AL. | 2020 | 9 MEN , 10 WOMEN, 1710 COD | HAWK EYE | DIRECTION CHANGE |
| MARTINEZ-GALLEGO ET AL. | 2021 | 24.982 VOLLEYS, 142 MATCHES | HAWK EYE | VOLLEY IN DOBLES |
| KOVALCHIK & ALBERT | 2022 | 142.803 POINTS, 1.334 MATCHES | TRACKING DATA | RETURN |
| VIVES ET AL. | 2023 | 14.146 1° SERVES | HAWK EYE | 1° SERVE IN DOUBLES |
| ZHOU ET AL. | 2024 | 36 PLAYERS, 42 MATCHES | KINOVEA | TYPES OF STANCES |

## CONCLUSIONS

The implementation of tracking data and new technologies in professional tennis has allowed for more detailed and in-depth analysis of the spatiotemporal characteristics of the game. This evolution has changed the way data is approached, leading to an increase in the number of studies employing machine learning (ML) or deep learning (DL) techniques. These trends indicate a significant shift in how performance in tennis is understood and analyzed, suggesting a promising future for the application of advanced analytical methods in this sport.

## DISCLOSURE STATEMENT

The author has not declared any potential conflicts of interest nor has received any funding to write this article.

## REFERENCES

Chase, C. (2020). The data revolution: Cloud computing, artificial intelligence, and machine learning in the future of sports. 21st century sports: How technologies will change sports in the digital age, 175-189.

Giles, B., Kovalchik, S. & Reid, M. (2020) A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis, Journal of Sports Sciences, 38:1, 106-113, DOI: 10.1080/02640414.2019.1684132

Kovalchik, S. & Reid, M. (2018) A shot taxonomy in the era of tracking data in professional tennis, Journal of Sports Sciences, 36:18, 2096-2104, DOI: 10.1080/02640414.2018.1438094

Kovalchik, S. A., & Albert, J. (2022). A statistical model of serve return impact patterns in professional tennis. arXiv preprint arXiv:2202.00583.

Martínez-Gallego, R., Ramón-Llin, J., & Crespo, M. (2021). A cluster analysis approach to profile men and women's volley positions in professional tennis matches (doubles). Sustainability, 13(11), 6370.

McCarthy, J. (2004). What is artificial intelligence?.

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F., Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big Hawk-Eye data. Statistical Analysis and Data Mining: The ASA Data Science Journal, 9, 310-325. https://doi.org/10.1002/sam.11316

Mitchell, T. (1997). Machine Learning. McGraw Hill. p. 2. ISBN 978-0-07-042807-2.

Russell, S. J., & Norvig, P. (2010). Artificial intelligence a modern approach. London.

Samuel A. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3):210-29.

Shimizu, T., Hachiuma, R., Saito, H., Yoshikawa, T., & Lee, C. (2019, October). Prediction of future shot direction using pose and position of tennis player. In Proceedings Proceedings of the 2nd International Workshop on Multimedia Content Analysis in Sports (pp. 59-66).

Stanko, I. (2020). The Architectures of Geoffrey Hinton. In: Skansi, S. (eds) Guide to Deep Learning Basics. Springer, Cham. https://doi.org/10.1007/978-3-030-37591-1_8

Stappen, L., Milling, M., Munst, V., Hoffmann, K., & Schuller, B. W. (2022). Predicting Sex and Stroke Success--Computer-aided Player Grunt Analysis in Tennis Matches. arXiv preprint arXiv:2202.09102.

Thakur, A., & Konde, A. (2021). Fundamentals of neural networks. International Journal for Research in Applied Science and Engineering Technology, 9(VIII), 407-426.

Vives, F., Lázaro, J., Guzmán, J. F., Martínez-Gallego, R., & Crespo, M. (2023). Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. Applied Sciences, 13(24), 13213.

Wei, X., Lucey, P., Morgan, S., Carr, P., Reid, M., & Sridharan, S. (2015, August). Predicting serves in tennis using style priors. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 2207-2215).

Whiteside, D. and Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach, Journal of Sports Sciences, 35:7, 648-654, https://doi.org/10.1080/02640414.2016.1183805

Zhou J. Q. & Liu, Y. (2024) Probability prediction of groundstroke stances among male professional tennis players using a tree-augmented Bayesian network, International Journal of Performance Analysis in Sport, DOI: 10.1080/24748668.2024.2314646