

Review Article

Time Series Prediction for Cycling Sport Performance Through a Machine Learning Predictor

**Muhammad Zulazri Hanis Mohd Nawi^{1*}, Sukhairi Sudin^{1,2},
Fathinul Syahir Ahmad Saad^{1,2}, Muhamad Khairul Ali Hassan^{1,2},
Faranadia Abdul Haris³, Nurul Syahirah Khalid^{1,2} and
Kamarulzaman Kamarudin^{1,2}**

¹*Faculty of Electrical Engineering and Technology, Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia*

²*Centre of Excellence for Advanced Sensor Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia*

³*School of Electrical Engineering, Universiti Teknologi Mara, Cawangan Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia*

ABSTRACT

Predicting sports performance has become a central focus in sports analytics, driven by the increasing availability of data and the growing recognition of its potential impact on decision-making in the sports sector. Time series analysis and real-time prediction of athletic performance involve forecasting an athlete's performance over time, allowing coaches and sports scientists to refine training programs, manage workload, and make informed strategic decisions. This study thoroughly examines time series prediction and real-time prediction in sports, as well as the artificial intelligence (AI) techniques employed by prior researchers. The review is conducted with precision,

strictly following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria. This article examines time series prediction and real-time prediction methodologies that utilize machine learning (ML) and deep learning (DL) approaches, spanning the period from 2020 to 2025. This article covers the range of AI methodologies from the most basic to the most advanced models. A detailed assessment of ML and DL methodologies, grounded in prior research findings, is presented. All approaches examined in this paper significantly influence the primary future study, which focuses on the hybrid long

ARTICLE INFO

Article history:

Received: 01 January 2025

Accepted: 29 April 2025

Published: 14 October 2025

DOI: <https://doi.org/10.47836/pjst.33.6.13>

E-mail addresses:

zulazrihanis7298@gmail.com (Muhammad Zulazri Hanis Mohd Nawi)

sukhairi@unimap.edu.my (Sukhairi Sudin)

fathinul@unimap.edu.my (Fathinul Syahir Ahmad Saad)

khairulhassan@unimap.edu.my (Muhamad Khairul Ali Hassan)

faranadia@uitm.edu.my (Faranadia Abdul Haris)

syahirakhkhalid@unimap.edu.my (Nurul Syahirah Khalid)

kamarulzaman@unimap.edu.my (Kamarulzaman Kamarudin)

* Corresponding author

short-term memory (LSTM) peephole integration with gated recurrent unit (GRU) for use in track cycling sports, the principal objective of the research. This research is consistent with the United Nations' Sustainable Development Goals.

Keywords: Cyclist performance prediction, hybrid predictor, performance classification, sports prediction, time series prediction

INTRODUCTION

Within the realm of athletics, there is a constant drive to further improve and strive for perfection. To achieve a competitive advantage, athletes, coaches, and analysts continually seek innovative approaches to enhance their performance. One example of this is the utilization of emerging technologies to acquire their skills more rapidly. The integration of AI is emerging as a transformative force in this dynamic landscape of sports performance enhancement. ML and DL are serving as the cornerstone of this transformation, with data-driven insights serving as the foundation. As a result of the application of technology in sports science to deliver real-time feedback, the way athletes receive input and improve their performance has undergone a revolutionary change in recent years. According to Vec et al. (2024), the exploitation of real-time data has revealed itself to be particularly beneficial when it comes to the provision of feedback to athletes.

The year 1997 marks the early application of AI in the sports industry. Prozone (Sports Business Journal [SBJ], 2023) was the first company to deploy AI in 1997 for the purpose of developing the first computer vision player tracking system in elite sports. This was the beginning of the field. A revolutionary computerized tracking system that can analyze movement patterns in a variety of sports was launched by Prozone (Di Salvo et al., 2006). Most of the company's efforts have been concentrated on quantifying not only the motion characteristics of professional football players but also the work rate ratios of these players during professional competitions. Not only does this method have the benefit of being applicable in official competitions, but it also has the potential to assist in the evaluation of high-level football performance.

According to SBJ (2023), the first officiating use case for Hawk-Eye computer monitoring in tennis was implemented in 2001. In addition to being known as the rapid reaction system, the Hawk-Eye system is a tool that helps referees quantify the accuracy of their penalties. Although the theory behind Hawk-Eye technology is quite straightforward, it is extremely accurate. It consists of eight to ten high-speed cameras, four computers, and a wide-screen display, and it determines whether tennis is played outside or indoors by analyzing the records of different ball positions. In some large-scale tennis courts, Hawk-Eye technology has evolved into an all-important tracking system. Tennis moves at a high speed during competitions; after landing, it is prone to generating issues and disputes about whether the tennis is in line or out of line, as well as other similar ones. Additionally, the

Hawk-Eye system can accurately record and judge ball landing point positions, thereby eliminating similar problems (L. Li & Shi, 2014).

One of the first known applications of AI in player personnel choices occurred in 2007, when Houston Astros consultant Ari Kaplan used it to make the case for Hunter Pence's promotion to the major leagues (SBJ, 2023). This was the first time AI had been used in this capacity. In 2009, AI was transformed into a new system that was given the name SportVU. This system is utilized in basketball sports and is applied through the ball. Additionally, a player monitoring camera system was presented to executives from the NBA. SportVU, developed by STATS, LLC, records data points for all ten players, the three referees, and the ball every thirty seconds. These data points measure speed, distance, player separation, and ball possession. According to an article written Richman (2018), the entire league would be equipped by the year 2013. It is planned to record, code, and categorize each movement, including every step, dribble, pass, shot, rebound, and other actions.

The year 2011 marked the year AI led the revolution in highlights through WSC Sports. The use of AI in this context involves the analysis of live sports broadcasts, which enables the identification of significant events in real time and the unlocking of limitless highlight combinations that can be easily distributed across many platforms. In the National Basketball Association (NBA), there are two stages of analysis. The first step is to conduct a real-time analysis of the game and determine what occurs during each second of the competition, from every goal to every touchdown. According to Lingeswaran (2023), each event is automatically detected and tagged with the precise day, time, and location at which it occurred. Additionally, they determine the beginning and ending of each play.

This evolution led Spiideo to introduce the first AI-powered, automated camera system for tracking sports in 2012. With the assistance of AI, it eliminates the friction associated with the roles of camera operator, producer, director, and commentator, thereby delivering production capabilities comparable to those of traditional broadcasting (Ruddock, 2024). A proliferation of chatbots from Satisfi Labs, Facebook Messenger, and Sapiens was observed in the sports industry in 2016. Satisfi Labs' first sports customer was the New York Mets (Anderson & Rainie, 2023). The Associated Press contracts Automated Insights to utilize its Wordsmith tool for game coverage on minor league baseball and high school sports that would otherwise go unreported (Mullin, 2016). Neither of these organizations would have been able to cover these types of sports otherwise.

Zone7 is the first sports science firm to use AI to anticipate the risk of athlete injuries. Their AI system can perform sophisticated computations on large amounts of data efficiently and consistently, thereby providing valuable insights daily. Due to its capability of performing these functions in a reliable manner in real-time, it reduces the amount of work required for data analysis. Additionally, it has the potential to make a significant contribution, even if it has not yet been fully utilized, to the way that professionals working

in multi-disciplinary teams (MDT) manage and minimize injury risks in professional athletes (Buchanan et al., 2022).

In 2019, the world was taken aback by the launch of a new technology that incorporates AI and is incorporated into the motion capture app, Physimax. This technology, which requires only a smartphone camera, plays a significant role in optimizing athletes' skills, techniques, and strategies by providing detailed feedback on motion data (Suo et al., 2024). A new regulation was implemented at the stadium, specifically for football matches, to reduce the risk of exposure to the coronavirus during the COVID-19 pandemic, which began at the end of 2019 and continued through 2020 and 2021. Express Access, powered by Wicket, will be introduced by the Browns in 2021 to provide fans with a more convenient and time-saving method for entering FirstEnergy Stadium. To provide participating fans with a touchless, faster, and more pleasant experience as they enter the stadium, Wicket's facial verification technology and computer vision analytics are utilized. According to Gribble (2021), this is the most recent initiative implemented as part of the Browns' Responsible Restart Plan. This plan was implemented in response to the COVID-19 outbreak. It will continue to be implemented to provide fans with the safest and most enjoyable experience possible when attending Browns games at FirstEnergy Stadium.

Semi-automated offside video-assisted review will be put through its paces at the 2021 Arab Cup, in preparation for the 2022 World Cup. During the Fédération Internationale de Football Association (FIFA) World Cup 2022, which begins on November 21 in Qatar, FIFA has announced that semi-automated offside technology will be used. This technology will serve as a support tool for both the video match officials and the on-field officials, assisting them in making offside decisions that are quicker, more accurate, and more reproducible on the most important stage of all (FIFA, 2022). There is also a partnership between WHOOP and OpenAI to produce WHOOP Coach, the first generative AI training assistant (WHOOP, 2023). In the meantime, International Business Machines Corporation (IBM) will introduce AI-generated video commentary for tournaments such as the Masters, Wimbledon, and the U.S. Open in 2023 (Hastings, 2023).

Sports Popular with AI

This historical occurrence, as documented above, demonstrates that AI technology can be incorporated into virtually every aspect of sports, including media, coaching, and player monitoring systems. This also includes almost every type of sport, such as cycling, football, tennis, baseball, or rugby, among others, that can be improved with the application of an AI system. This can be done to enhance aspects such as player performance, ball trajectory, or clear view in referee systems, such as the predominant system of positioning and division of responsibility used by football match officials worldwide. This section examines three popular sports that have implemented AI technology. The dynamic nature

of the sport itself, which necessitates consideration of various circumstances, contributed to the selection of this sport. Football, basketball, and cycling are the sports that will be examined in this article.

Several studies have already been conducted in the field of football or by earlier researchers, such as Aleza and Vetrihangam (2023), which apply AI to prevent errors in the refereeing process of a football match. This is because any error or poor decision made by the referee during a match has the potential to completely change the course of the match and result in heartbreaking defeats. As a result, the findings of their research center focus on examining ways to improve the decision-making process in football, while also highlighting the limitations of the process in cases that require interpretation. They concluded that AI could revolutionize the decision-making process not only in football but also in other sports. This is something that they discovered via their investigation. While this is happening, other researchers, such as H. Li et al. (2024), are implementing AI in football to study methods that may improve the quality of football instruction through AI and metaverse applications in mobile Internet environments. By combining the metaverse and internet or Internet of Things (IoT) applications, they successfully translated the development of football teaching quality into a research project on a 360-degree panoramic football teaching video distribution method. Additionally, researchers such as Gomes et al. (2013) are investigating the development of an Intelligent Decision Support System to predict the outcomes of football games, with a particular emphasis on betting procedures.

As is the case in football, playing basketball also involves coaching to identify areas where talented new players can be developed. It is therefore possible to improve the process of teaching and training sessions by utilizing the AI system. Additionally, the new player will only require a shorter amount of time to comprehend the material as compared to the conventional technique. An earlier researcher, Tingting (2015), has provided the collegiate basketball team with an Intelligent Teaching System based on the development of an AI system. Her teaching method is a multi-level, dynamic, interactive multimedia presentation system that possesses good interactivity and intelligence. She presents this system to the classroom. Teachers and students can be based on the requirements of teaching and learning, as well as the selection of effective learning tools, if a logical relationship is established between the information contained in multimedia. The practical test demonstrates that her research has the potential to increase the effectiveness of basketball instruction in physical education colleges and universities, and it also has a certain promotional value. Her study has accomplished the design goals and requirements that were anticipated. Other researchers, such as T. Xu and Tang (2021), have utilized an AI system in basketball to prevent sports injuries caused by collisions in the sport. The researchers conducted an in-depth investigation of the shooting action, as well as an analysis of the fundamental training activities and the sport recognition involved in basketball training. During this

interim period, Yan et al. (2023) have successfully implemented an AI system in basketball by introducing basketball shooting analysis. This is because it has the potential to help basketball players refine their shooting methods and accuracy, which in turn can enhance the efficiency of both games and training sessions.

Cycling, which is separated into numerous categories such as track cycling, road bicycle racing, cyclo-cross, mountain bike racing, freestyle bicycle motocross (BMX), and cycle speedway, is a different sport from the two sports mentioned above (Wood, 2020). Balance, coordination, and riding tactics are some of the fundamental skills that are needed for all cyclists to hone (Proulx & Smith, 2014). Although cycling sports are classified into a few categories, they all share a common core talent that is essential for all cyclists to develop. In more advanced fields, it may be necessary to possess extra skills, but the fundamental abilities remain the same. Using AI, Ahmadi et al. (2024) were able to accurately estimate radial and mediolateral force components while cycling by utilizing inexpensive sensor data and participant characteristics. This was made possible by research conducted by Ahmadi et al. (2024). The reason for this is that precise measurements of pedaling kinetics and kinematics are crucial for maximizing the effectiveness of rehabilitation, exercise training, and understanding musculoskeletal biomechanics. There is a strong correlation between the pedal reaction force, which is the primary external force in cycling, and the activity of the lower limb muscles, as well as the joint reaction forces. This force is essential for musculoskeletal models.

An alternative perspective is presented by researchers such as Wu et al. (2020) regarding the use of intelligent cycling equipment to establish personalized training intensities for individual cyclists. In the meantime, Ibrahim et al. (2017) reported that their team was utilizing AI to forecast the chance of bicycle near misses. This was done to avoid potentially hazardous situations that could result in severe consequences, particularly during competitions or races held in urban areas. Different from Karetnikov (2019), who used the application of AI into cycling to analyze for professional cyclists and propose a new framework that aims to provide the coach with a method with which he or she can predict the performance in an upcoming race based on recently completed training sessions and planned training sessions in the near future, this framework aims to provide the coach with a method that predicts the performance of the athlete in an upcoming race. In this manner, the coach can evaluate several planned alternatives to determine the most effective method of preparing the athlete for competition. Lastly, but certainly not least, N. Li et al. (2024) and their fellow researchers utilized AI to investigate whether it can accurately predict the maximum volume of oxygen (VO_2) that the body can utilize during exercise across various exercise intensities by combining muscle oxygen (MO_2) with heart rate (HR). They used twenty young athletes who had received extensive training to conduct the following tests: a ramp incremental exercise, three submaximal constant-intensity exercises, and three severe, exhaustive-intensity exercises.

The Affected Factors on Sports Player Performance

As a result of deploying the AI system described above, it is evident that numerous aspects influence the sports player. Whether it be cyclists, football players, or basketball players, all these factors have the same component. In this section, the factors that affected player performance will be explained, including physical, cardiovascular, psychological, and any other relevant elements. Athletic success and elite performance are frequently the product of the best possible fusion and coordination of all aspects of sports readiness (i.e., physical fitness, tactical and technical expertise, bioenergetic and neuromuscular capacities and abilities, anthropometric traits, biomechanics, physiological, emotional intelligence, creativity, or personality), which develop because of methodical, sustained sports preparation. In contrast, individual and team sports have different weights assigned to these attributes (Calleja-González et al., 2023). While certain individual sports demand a high degree of bioenergetic, neuromuscular, and other capacities, the technical and

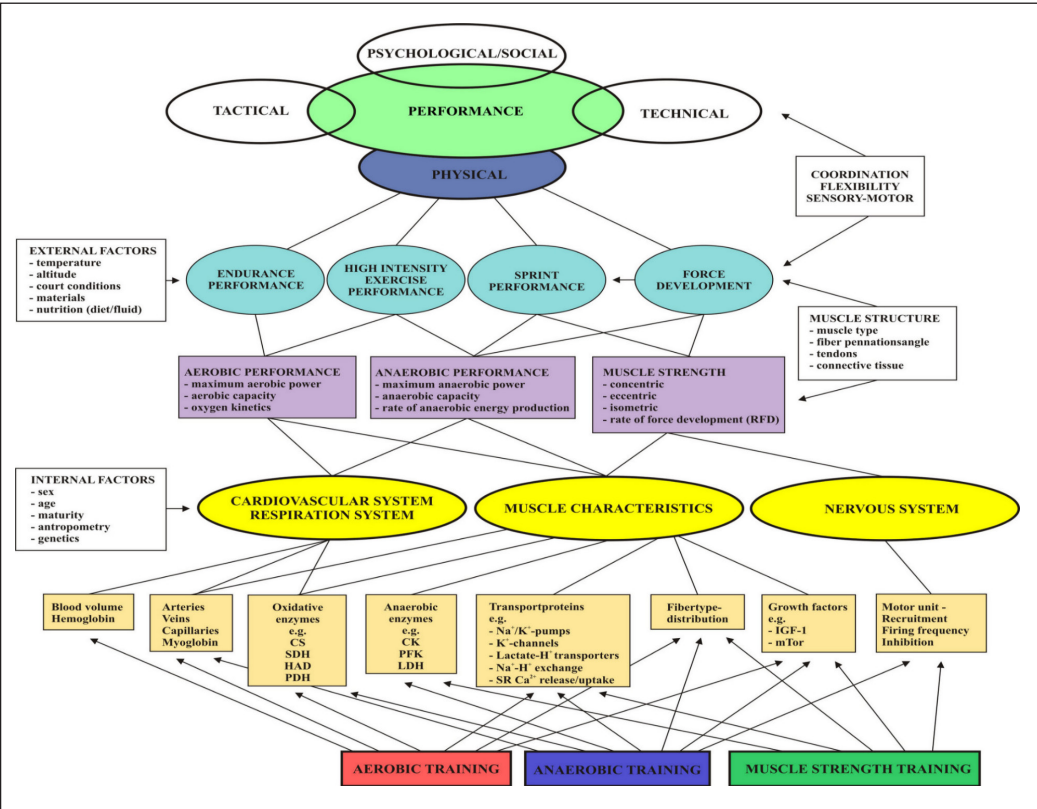


Figure 1. Comprehensive model of the factors that affect sports performance (Bangsbo, 2015)
Note. CS = Citrate synthase; SDH = Succinate dehydrogenase; HAD = Histamine dehydrogenase; PDH = Pyruvate dehydrogenase; CK = Creatine kinase; PFK = Phosphofructokinase; LDH = Lactate dehydrogenase; Na^+ = Sodium ion; K^+ = Potassium ion; H^+ = Hydrogen ion; SR = Sarcoplasmic reticulum; Ca^{2+} = Calcium ion; IGF-1 = Insulin-like growth factor 1; mTor = Mechanistic/mammalian target of rapamycin

tactical skills of team sports are intimately linked to performance and can make up for deficiencies in fitness levels (Bangsbo, 2015; Dolci et al., 2020). Bangsbo (2015) created a comprehensive model of the factors that affect sports performance in Figure 1.

Gavin (2022) asserted that a person's ability to succeed in sports is influenced by their sleep patterns, emotional condition (stress, anxiety, and sadness), pre-match routine, and romantic relationships. During his study on the factor of player performance, he did not take a bias towards any sport; instead, he exclusively focused on the factor.

According to the Global Web Index (GWI), as noted by Seedhouse (2017), since football is one of the world's major sports, various factors influence player performance based on their style of play. The factors encompass physical, psychological, tactical, and environmental aspects, all of which impact a football player's performance. Physical aspects include exhaustion and fitness levels. While physical fatigue (Dambroz et al., 2022) can negatively impact technical skills such as shooting, dribbling, and passing, high levels of both aerobic and anaerobic fitness are essential for maintaining performance throughout the game (Plakias & Michailidis, 2024). Players who possess higher levels of motivation and confidence tend to outperform others (Nakisa & Rahbardar, 2021), whereas those who experience high levels of stress and anxiety may struggle to make decisions and perform well overall. These psychological elements include motivation and worries (Nakisa & Rahbardar, 2021).

The two main tactical aspects are opposing quality and team strategy. A player's performance can be impacted by the quality of the opposition team, particularly in high-pressure situations, and the team's tactical approach, including formations and playing style, can affect individual performance (Bangsbo, 2014). The playing surface and weather conditions make up the environmental aspects. Extreme weather, such as heat or rain, can impair endurance and physical performance, and different types of playing surfaces, like grass or artificial turf, can affect performance and injury risk (Bangsbo, 2014). Coaches and athletes can enhance their training and match preparation to improve performance on the field by gaining a deeper understanding of these variables.

A basketball player's performance can be significantly influenced by several factors, including their level of physical fitness, psychological well-being, core training, performance analytics, shooting accuracy and technique, game circumstances, and external variables. It is essential to possess physical fitness qualities such as strength, speed, endurance, agility, and flexibility. A player's capacity to sprint, jump, balance, and change direction rapidly is influenced by these physical characteristics (S. Luo et al., 2023). Basketball players may be impacted by psychological issues since stress reduction, mental toughness, and attention are essential. Players who can remain composed under duress typically play better (S. Luo et al., 2023). Basketball players must engage in core training because total athletic success depends on having a strong core. Core training has been linked to improvements

in strength, agility, balance, passing, dribbling, shooting, and rebounding, according to studies (S. Luo et al., 2023).

Lastly, external circumstances and game conditions can significantly impact a player's performance throughout a match. Coaching techniques, crowd dynamics, and game location can all impact a player's performance. The comfort and confidence levels of a player might be impacted, for example, by playing at home vs. away (F. Wang & Zheng, 2022). The sum of these elements affects a player's overall performance on the court.

Between these sports, the factors that affect cyclists can be seen as having a significantly more profound impact on them if these factors occur. As for cyclists, the physiological, biomechanical, nutritional, aerodynamic, and physical components of elite cycling performance have all been examined (Atkinson et al., 2003; Craig & Norton, 2001; Faria et al., 2005a, 2005b). As further explanation, the individual cyclist factors, such as physiological attributes, such as aerobic capacity, muscle fiber composition, and lactate threshold (Phillips & Hopkins, 2020). Meanwhile, nutrition and hydration, as a proper diet and hydration, are crucial for maintaining energy levels and endurance (Phillips & Hopkins, 2020). Training and conditioning cannot be neglected, as these can contribute to individual cyclist factors that regular and targeted training can significantly enhance the performance of cyclists (Faria et al., 2005b).

To optimize a cyclist's performance during a match, tactical considerations are also necessary in cycling. Two examples of tactical variables in cycling sports are race strategy and interpersonal dynamics. Choosing the right pace, drafting position, and moment to put in your best effort all affect how well you perform; therefore, race strategy is crucial (Faria et al., 2005b). Teamwork, competitiveness, and interpersonal interactions with other cyclists are factors (Phillips & Hopkins, 2020). Wind, temperature, and humidity are just a few examples of the environmental elements that can have a significant impact on a player's performance during a match (Phillips & Hopkins, 2020). The effort required is also influenced by the kind of terrain, such as flat, hilly, or mountainous (Faria et al., 2005b).

Gear ratios, cadence, and bicycle layout are examples of equipment factors that can provide a unique match experience for every competitor. Bicycle setup, which encompasses both mechanical effectiveness and aerodynamics, is crucial in creating a distinct margin between players (Faria et al., 2005b). Pedaling cadence can increase efficiency throughout a match, and gear ratios and cadence can be used to adjust the ideal gear selection (Faria et al., 2005b). In addition, and just as importantly, socio-cultural and economic aspects of the world can influence the ease of resource acquisition, the strength of supportive networks, and the impact on possibilities for training and overall performance (Phillips & Hopkins, 2020).

To sum up, Lastella et al. (2015) stated that sleep and recovery are also one of the major patron preparations for cyclists, as cyclists with adequate sleep and recovery can

achieve better performance enhancement. It is possible for teams, coaches, and cyclists to create more successful training and competition plans when these aspects are understood.

The objective of this study is to investigate the current state-of-the-art in time series prediction and real-time prediction applications in sports, with a particular focus on cycling sports. However, this review will also encompass other sports, such as basketball and football/soccer, as well as research accomplishments, and it will recommend prospective future paths. A review of prior research on time series prediction in sports applications, primarily from the years 2020 to 2025, is presented in this publication. The purpose of this research was to examine the development of techniques over this time in chronological order. These techniques included ML and DL approaches, which were particularly concerned with real-time prediction as opposed to batch prediction when applied to datasets. This presentation is valuable since it is geared towards professionals in the sports industry, including sports scientists, coaches, and, most importantly, the players themselves, to motivate them to self-analyze overtime and maintain their performance in competition. The main contributions highlighted in this paper are:

- An in-depth investigation into how ML and DL models are applied in the realm of sports, as well as their respective methodologies.
- A methodical investigation of the application of time series prediction and real-time prediction, which is combined with ML and DL techniques and is utilized in the sports industry.
- The results of previous researchers on applying ML and DL in sports will be analyzed, and an evaluation of the field of applying these technologies in sports will be taken into consideration.
- Comparisons and discussions have been made regarding the most recent state-of-the-art procedures.

PROBLEM STATEMENTS

The primary objective of a more extensive body of study, which includes this review, is to utilize contemporary technologies to expedite the learning process in professional sports. It was estimated that the annual revenue of the professional sports sector is currently 159 billion dollars (Sim, 2024). This makes it a massive enterprise. This suggests that there is a significant possibility for any new technology that could provide the team using it with even the slightest advantage over its competitors. It is possible to acquire this competitive advantage by providing athletes with time series prediction feedback if they make a mistake, or by checking their performance while they are competing in sporting events (Hribernik & Kos, 2023). Currently, one of the most pressing questions on our minds is why the world needs AI, irrespective of whether it is ML or DL, in the sports business. In the sports business, human errors and faulty decision-making in football matches are two

of the issues that need to be addressed by AI (Aleza and Vetrihangam, 2023). A mistake like this could result in the referee making an incorrect decision during the match.

Additionally, according to the findings of other researchers, the teaching approach in physical education at universities continues to adhere to traditional teaching methodologies, particularly in the realm of sports, where concepts that are integral to the teaching work and should be enhanced accordingly are often overlooked (H. Li et al., 2024). The occurrence of sports injuries is an unavoidable issue for athletes participating in any sport, and it is also a circumstance that cannot be avoided during training and competition. The high-intensity and high-speed collisions that occur between players in basketball are one of the reasons why basketball is considered one of the sports that pose the greatest danger of injury. Because of this, we require the involvement of AI to either prevent it or, at the very least, mitigate the risk it poses to the player (T. Xu & Tang, 2021). In the meantime, the decision regarding the training plans that will lead to the greatest performance in the competitions is mostly based on the specialist knowledge of the team coach in the sport of cycling. Several recent advancements in the fields of sports engineering and data science have enabled the exploration of new avenues for data gathering and analysis driven by data. To support these decisions, it suggests new prospects for the use of AI, which is why the assistance from AI is required (Karetnikov, 2019).

However, before beginning to use AI, especially ML or DL, for event time series prediction to deliver feedback, the status of the subject must be investigated. The primary purpose of this study is to critically evaluate and examine works published between 2020 and 2025 that focus on the use of ML and DL algorithms to provide players with time series prediction and real-time feedback for any sport.

DATABASES, KEYWORDS, AND INCLUSION CRITERIA

In this paper, the PRISMA framework (Page et al., 2021) is followed for reviews. Three databases were searched, including Scopus and Web of Science to cover interdisciplinary research, as well as IEEE Xplore, which covers articles, conference proceedings, and other publications. The first database search was conducted in August 2024, during which relevant journals and conference proceedings were identified and extracted. To update the analysis, an additional search was conducted in October 2024 and February 2025 to include journal articles from December 2024 to February 2025. This paper aims to review all articles in the last six years that focus on using ML and DL to provide real-time feedback and time series prediction to athletes and to whatever sports they play, but mostly focused on cycling, and include other sports as well, like football and basketball, where they track the athletes using whatever equipment and sensors to capture. The following keyword string was used for Web of Science, IEEE Xplore, and Scopus: (“machine learning” OR “deep learning”) AND (real time OR “real-time” OR real-time OR concurrent OR instant) AND (sport OR

exercises) AND (time series OR prediction OR predictor OR predict) (time AND series AND prediction AND in AND sport).

Articles that did not provide a detailed description of the procedure were not taken into consideration. Additionally, publications that focused on the generalization of activities into major tasks, such as lying, sitting, or walking, were excluded from the compilation. Additionally, articles written in languages other than English were excluded.

A summary of the inclusion and exclusion criteria is presented in Table 1. As a result of the fact that several of the papers satisfied more than one of the criteria for exclusion, the total is higher than the total number of publications that were not included.

The scoping review primarily assessed real-time or time series prediction methodologies utilized in the current literature and investigated the accompanying research questions:

1. Data: What data sources were utilized? What was the magnitude of the sample size?
2. Preparation: Was dimensionality reduction implemented? What method was employed to partition the data set into training and validation subsets?
3. Methodology: What methodologies were employed? Do these methods consider the interdependencies among the segments or individual movements?
4. Validation: What measures were employed for performance assessment?

Study Selection

Keyword searches of the three databases retrieved 22,684 articles covering the period from August 2024, along with an additional 269 articles that included journal articles from December 2024 to February 2025. After screening the titles and abstracts, 3,250 articles remained for detailed full-text analysis. Of these 3,250 articles, 75 were relevant for inclusion. For quality assurance, each full text was screened, and data extraction was performed by using the Mendeley tool to extract the information systematically. An amount of 1,540 articles were not related to sports or addressed a different application area, 1,230 articles were excluded because the time series task was not included at all or was not one of the focuses of the article, 150 articles did not clearly describe the analysis or were not found, 125 articles that miss out from title and abstract screening and 110 articles were conference papers that were preliminary work for an included journal article that included the same methods or an extension of an existing work without any new method. Finally, 15 papers were excluded because their full texts were not in English. Figure 2 provides an overview of the inclusion process.

Table 1
Inclusion and exclusion criteria

Inclusion	Exclusion
2020–2025	Outside 2020–2025
English	Full text not in English
The method is clearly described	No clear description
Active human movement	Not human movement
Time series prediction or real-time	Other than time series prediction or real-time
Specific sports activity	General activity

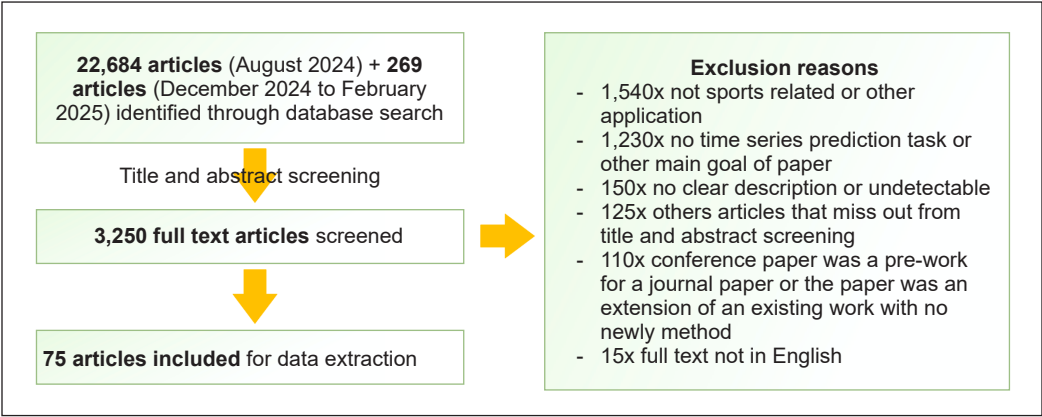


Figure 2. Flow diagram of the review

Recent Development of the Field

The field of using ML and DL in sports is developing rapidly. This paper splits the included articles by whether they use ML or DL. The reviewed articles indicate that performance analysis may be a vital tool to help coaches make decisions at the highest levels (Wright et al., 2013). There are many different prediction techniques, the two most common being linear and nonlinear (Zhang et al., 2022). Among these, nonlinear prediction should be used to forecast sports performance since a variety of factors influence sporting events. The physical and psychological state of the participants, the strategies used in the game, biomechanical, and even environmental aspects might have an impact on the outcome (Elmagd, 2019). As a result, although the nonlinear approach may take these elements into account, the linear prediction technique cannot work the same.

Nowadays, to improve decision-making and obtain insights into sports performance, sports organizations are depending more on advanced prediction models (Cojocariu, 2022). Traditional analytical techniques have given way to more dynamic, data-driven methods, which allow for a complete understanding of the complex aspects affecting sports results (Fujii, 2021). When working with time series data, data analytics becomes very important in the sports industry. The forecasting of time series athletic performance has been a crucial focus, realizing the necessity of identifying patterns and trends in sequential data (Kolambe & Arora, 2024). There is little doubt that predicting performance is difficult, and predicting long-term performance is much more difficult, with little research done on the subject to date (Pantzalis & Tjortjis, 2020).

Hence, below we split the application of time series prediction into two categories: time series prediction and real-time prediction using ML techniques, and time series prediction and real-time prediction using DL techniques. The taxonomy of ML and DL, as shown in Figures 3 and 4, helps to distinguish between techniques in ML and those in DL.

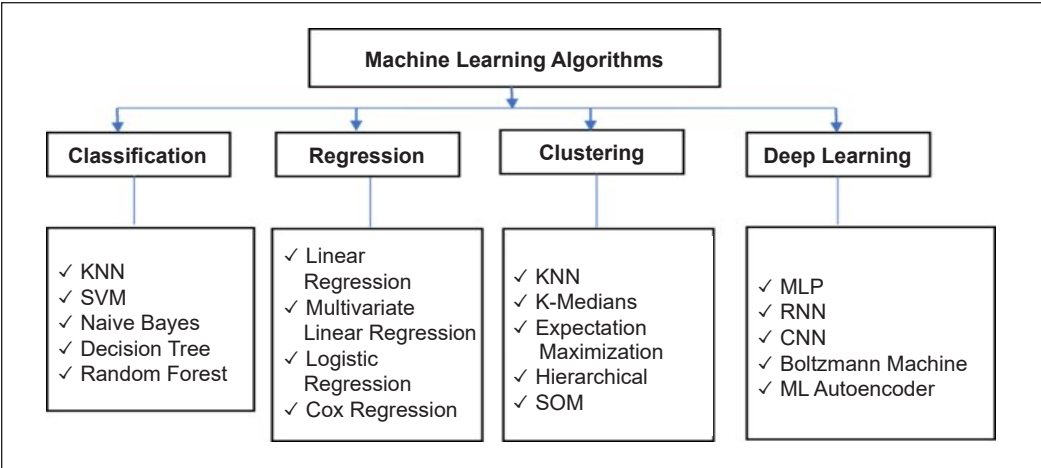


Figure 3. Taxonomy of machine learning algorithms (G. Wang et al., 2020)

Note. KNN = k-nearest neighbors; SVM = Support vector machine; SOM = Self-organizing map; MLP = Multilayer perceptron; RNN = Recurrent neural network; CNN = Convolutional neural network; ML = Machine learning

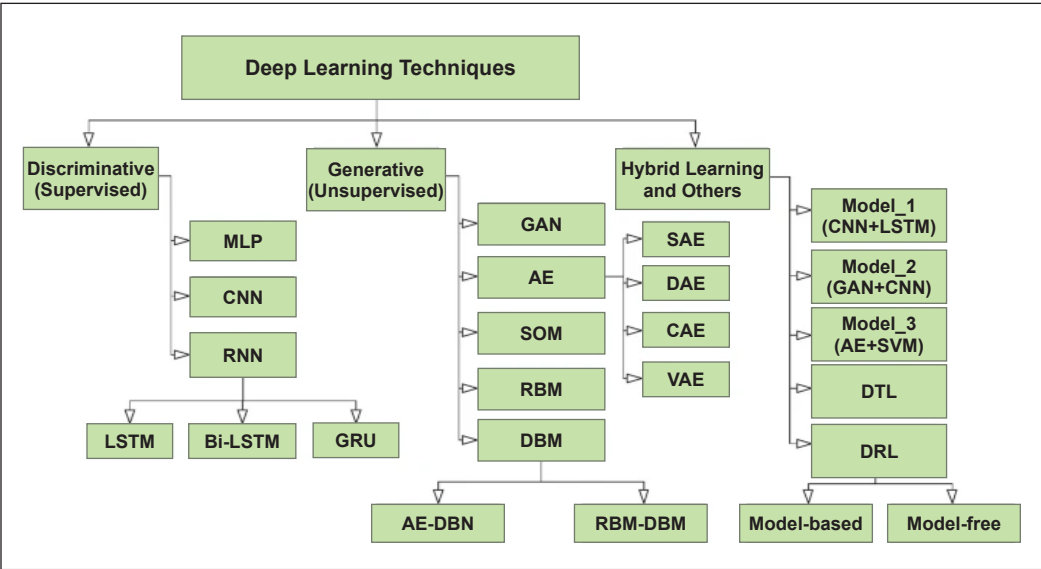


Figure 4. A taxonomy of deep learning techniques by Sarker (2021) was broadly divided into three major categories: (i) deep networks for supervised or discriminative learning, (ii) deep networks for unsupervised or generative learning, and (iii) deep networks for hybrid learning and others

Note. MLP = Multilayer perceptron; CNN = Convolutional neural network; RNN = Recurrent neural network; LSTM = Long short-term memory; Bi-LSTM = Bidirectional long short-term memory; GRU = Gated recurrent unit; GAN = Generative adversarial network; AE = Autoencoder; SAE = Sparse autoencoders; DAE = Denoising autoencoders; CAE = Computer-aided engineering; VAE = Variational autoencoders; SOM = Self-organizing map; RBM = Restricted Boltzmann machine; DBN = Deep belief network; AE-DBN = Autoencoder-Deep belief network; RBM-DBN = Restricted Boltzmann machine-Deep belief network; DTL = Decision tree learning; DRL = Deep reinforcement learning

ML METHOD OF TIME SERIES AND REAL-TIME PREDICTION

Linear Regression

Linear regression analysis is used to forecast the value of one variable depending on the value of another variable. The variable that the user intends to forecast is referred to as the dependent variable. The variable employed by the user to forecast the value of another variable is referred to as the independent variable (IBM, 2021). This technique was employed by the researcher (Yuan, 2024) to forecast the outcomes of certain basketball athletes involved in NBA competitions. The outcome of employing this technique indicates that linear regression emerges as the most effective model among the other models they tested for forecasting the performance of specific basketball players. The outstanding performance of the linear regression model, evidenced by its minimum mean squared error (MSE) of 0.0, implies an impeccable fit to the training data. Other researchers who have utilized linear regression analysis in their studies include De Beéck et al. (2018). Their research aims to forecast the rating of perceived exertion (RPE), a validated subjective metric of fatigue, utilizing inertial sensor data from individuals running outside. They attained a mean absolute error (MAE) ranging from 1.89 to 2.02.

Multiple Linear Regression

There was a study in 2020 that employed multiple linear regression (another variant of linear regression method) to forecast football player performance with the dataset consisted of 59 central defenders having played at least 10 matches for the English Premier League 2016–17 season and the result they obtained was 0.907 R^2 and 0.88 adjusted R^2 with the thirteen features were proved to be statistically significant (Pantzalis & Tjortjis, 2020).

Least Absolute Shrinkage and Selection Operator (LASSO) Regression Model

Another form of this structure (linear regression) is termed the LASSO regression model, which incorporates a penalty term to deter the coefficients of the independent variables from becoming excessively large (Wohlwend, 2023). The researchers who employed this technique are Shukla et al. (2022), which is utilized to forecast the overall revenue earned by a sports facility. The lasso regression model provides the most accurate forecast of total income earned by any sports facility on a specific day of the month, with an R^2 score of 0.97965, a MAE of 63.66423, and a root mean square error (RMSE) of 82.02202.

Generalized Linear Model (GLM)

Aarons et al. (2023) employed the GLM technique, a versatile extension of ordinary linear regression (Pennsylvania State University [PSU], n.d.), to predict the end-of-match outcomes of Australian football matches in real-time. Their classification accuracy was

initially low at the start of the match (45.7–48.8%) but improved significantly, reaching a peak near the match's conclusion (87.2–92.7%).

Support Vector Machine (SVM)

SVM is a supervised ML technique employed for classification and regression tasks. Although capable of addressing regression issues, SVM is especially adept at classification tasks (GeeksforGeeks, 2025d). Cui et al. (2023), demonstrated an application of this technique in wearable devices to analyze data related to body recovery and injury prevention in physical education. This approach facilitates the real-time collection of students' exercise data, encompassing metrics such as exercise volume, heart rate, steps, and distance using wearable technology. Their functional movement screening (FMS) test results ranged from 0.0756 to 0.0838, depending on the assessed activity.

From another perspective, Zheng and Yuan (2022) employed the SVM motion analysis model technique to analyze the time series data about the prediction and features of dance sport movements, achieving a recognition accuracy of 97 to 98%. Besides that, Hu (2024) utilized it for sports performance prediction through time series analysis of the physical performance of college students. He asserts that this prediction holds significant reference value for the university sports management department in planning and developing physical education courses, aligning the curriculum more closely with the actual needs of students and thereby establishing a more rational and scientific framework for college student physical training and daily exercise. The outcome of his proposed method, determined by the relative inaccuracy at each stage, ranged from 1.56 to a minimum of 0.06%.

Ren et al. (2024) also employed this strategy in their study on a real-time injury monitoring system. The motion depth learning technique was employed in their paper to identify activities of human motion targets. A real-time monitoring system for sports injuries was developed, encompassing the monitoring of human physiological data and the detection of target actions in human motion. The outcome achieved was 94.2% precision and 92.5% sensitivity.

Gradient Boosting

Gradient boosting is a widely utilized boosting approach in ML for classification and regression tasks. Boosting is a type of ensemble learning strategy that trains models progressively, with each subsequent model aiming to rectify the errors of its predecessor (GeeksforGeeks, 2025c). It amalgamates multiple weak learners to form robust learners. Yuan (2024) employed this technique during testing with linear regression, yielding a relatively low MSE of 0.05 for the gradient boosting model. Although it does not match the accuracy of linear regression, it still exhibits a significant level of precision in forecasting player performance in basketball.

eXtreme Gradient Boosting (XGBoost)

XGBoost is a scalable and distributed form of the gradient boosting framework, specifically a gradient-boosted decision tree (GBDT) ML library (NVIDIA Corporation, n.d.). Miah et al. (2022) employed this strategy in their study to anticipate and provide recommendations for both small and large sports teams regarding the suitability of an athlete for a specific game, utilizing a mobile application and mobile health data (m-health). The m-health utilizes mobile devices to collect clinical health information and monitor patients' vital signs in real-time. XGBoost outperformed the other studied ML algorithms, achieving an accuracy of 95.2%, a sensitivity of 99.5%, a specificity of 99.5%, and an F1 score of 99.66%. Zhu (2024) also employed the XGBoost approach to concentrate on real-time predictions of NBA game results. According to his findings, XGBoost models attained the maximum accuracy, roughly 99.88%. These results demonstrate their enhanced efficacy in forecasting NBA game results. He asserts that this will provide substantial promise for the integration of high-dimensional statistics and DL in sports analytics.

k-nearest Neighbors (KNN)

KNN is a non-parametric algorithm, a supervised learning classifier that uses proximity to classify or predict the categorization of a certain data point. It is a widely used and straightforward classifier for classification and regression in contemporary ML (IBM, n.d.b). Due to the enthusiasm surrounding the topic, Amendolara et al. (2023) conducted a comprehensive review of prior studies that employed this strategy to predict sports injuries. Their study states that KNN has practical restrictions regarding the sample sizes it can efficiently analyze. Nonetheless, its simplicity and versatility are evident. The incorporation of specialized sensors for enhanced data collection has refined KNN injury identification models and augmented their capacity to discern elements contributing to damage. The improved identification of predictive injury characteristics at the individual athlete level enables coaches and medical staff to modify training techniques to mitigate the recognized risk of injury.

Bayesian Inference

Bayesian analysis, a statistical inference method named after English mathematician Thomas Bayes, enables the integration of previous knowledge of a population parameter with evidence derived from sample data to inform the statistical inference process. A prior probability distribution for the parameter of interest is established initially. The evidence is subsequently acquired and integrated using Bayes' theorem to yield a posterior probability distribution for the parameter. The posterior distribution serves as the foundation for statistical judgments regarding the parameter (Britannica, 2025). In short, Bayesian inference is a method of statistical inference wherein the statistician gives subjective probability to the potential distributions that may produce the data (Taboga, 2021).

Autoregressive Integrated Moving Average (ARIMA)

ARIMA is a distinct methodology for forecasting time series data. ARIMA synthesizes these two methodologies, therefore its designation. Forecasting is a domain of ML that utilizes the historical behavior of a time series to predict one or more future values of that series (Noble, 2024). Almaleck et al. (2024) optimized this technique in their research to predict the 24-hour-ahead electrical usage of sports venues. The primary experimental implementation was conducted at Odd Football Club's Skagerrak Arena in Norway, yielding results of 15.29% for mean absolute percentage error (MAPE) and 30.61% for MAE, both of which are somewhat elevated. However, upon integrating the basic ensemble method (BEM) of neural networks with ARIMA to formulate a hybrid model, the value alters significantly to 9.03% and 20.24%. While the research conducted by his team appears to diverge from others by focusing on the electricity consumption of sports events rather than predicting sports performance, the methodologies employed by his team may be evaluated and utilized in our future research endeavors.

Random Forest

Random forest is a widely utilized ML technique, which amalgamates the outputs of numerous decision trees to produce a singular outcome. The simplicity and adaptability of the system have driven its widespread use, as it effectively addresses both classification and regression challenges (IBM, n.d.a). Stübinger et al. (2020) employed this technique in their research, comparing various models to predict the outcomes of football matches based on match and player attributes. Among all the models evaluated by their team, the random forest technique had the highest accuracy of 81.26%, the lowest RMSE, and the least mean absolute deviation (MAD).

Others

Other researchers created their own model of AI or used methods that were not specified. The researchers, like N. Li et al. (2024), developed an ML model to predict volume of oxygen (VO_2) at various exercise intensities by integrating muscle oxygen (MO_2) and HR in cycling. Their findings indicated that all models yielded comparable results, although the accuracy of VO_2 predictions varied among the models at different exercise intensities. Notably, the left model (LM) + right model (RM) + heart rate (HR) model exhibited a robust correlation, producing an r value of 0.94 with a significance level of $p < 0.001$.

Other researchers, Eid et al. (2024), also developed an ML hybrid model that illustrates the effectiveness of the Sport AI Model (SAIM) framework in improving the precision of sports-related outcome predictions. They evaluated their models to analyze English Football League sports data, aiming to gain in-depth insights into player performance in the English Football League. They assert that the incorporation of big data analytics, facilitated

by the SAIM model, can transform football strategies, enhance player performance, and promote data-driven decision-making, noting that their model has successfully predicted player performance in real-time.

There is also a unique method that was applied in their research from Almarashi et al. (2024), who employed a distinctive methodology in their research, utilizing a dynamic technique that examines performance structure through both linear and nonlinear time series models, specifically the non-linear autoregressive neural network (ARNN) model, alongside traditional stochastic linear and nonlinear models such as ARIMA, Exponential Smoothing (ETS), and Trigonometric Seasonal Box-Cox autoregressive moving average (ARMA) Time Series (TBATS). Their research focuses on developing predictive models for performance metrics that can aid both tennis players and sports analysts in predicting player rankings in upcoming matches. Their findings indicate that the ARNN model surpasses all other competing models, as seen by lower values of RMSE, MAE, and MAPE. The prediction results derived from the ARNN model indicate low 95% confidence intervals, signifying enhanced accuracy and dependability in the forecasts.

Table 2 presents the best-performing methods from across all methods used in ML to predict sports-related activities, incorporating time series prediction, real-time prediction, or capturing temporal dependencies.

Summary

The entire body of research publications and journals that dealt with the application of ML techniques to the prediction of time series or real-time predictions for sports applications was examined. The period covered by the paper that was examined is between 2020 and 2025. A thorough examination of the application in sports, the performance of the AI models, and the outcomes that they achieved were the primary considerations that led to the selection of the findings from the comparison table. Each of the findings from the comparison table will also serve as a guide for the subsequent study, which will be conducted in this article, focusing on the construction of the Peephole-GRU Predictor for Cyclist Player performance prediction. A further observation that can be made is that between the years 2020 and 2025, there was a minimal number of researchers who conducted time series prediction using the ML method. Furthermore, a small number of researchers successfully validated the AI model in practical situations, particularly in sports applications. Study by N. Li et al. (2024) became one of the main references (apart from the DL method) for our research. This is because their method utilizes the LM + RM + HR model, which they have created and modified from the vanilla ML method, to predict VO_2 across different exercise intensities by combining MO_2 with HR in cycling sports. This method was chosen from among the top five models in the comparison table. The reason for this is not only that the model they have adapted is highly suited and compatible

Table 2
Comparison of the best methods (best of five) by previous researchers based on results obtained for implementing time series prediction and real-time prediction in machine learning methods

References	Algorithm method	Result		
		Model concept	Accuracy	Limitation
Zheng and Yuan (2022)	SVM	Vanilla	97–98% for recognition accuracy	Since the structure used does not have a higher level of supervision, the effect is not ideal in some aspects
Aarons et al. (2023)	GLM	Upgrade the variant of linear regression	Classification accuracy was low at the start of a match (45.7–48.8%) but increased to a peak near the end of a match (87.2–92.7%)	Only considers technical performance through technical Pls; other aspects of performance are not accounted for. Also, the predictions are based on the performances of all teams in the AFL, which do not provide insight into the specific game styles of individual teams
Zhu (2024)	XGBoost	Upgrade the variant of gradient boosting	Achieved the highest accuracy, reaching approximately 99.88% compared to another model	Required substantial computational resources, challenges in interpretability, and necessitating additional tools to understand the decision-making process; also, insufficient or noisy data could hinder the model's ability to learn meaningful patterns
N. Li et al. (2024)	LM + RM + HR model	Unspecified ML Model	Very strong correlation that produce ($r = 0.94, p < 0.001$)	Complexity and time-consuming to train and implement the model
Almarashi et al. (2024)	ARNN	Hybrid	95% confidence intervals, indicating higher accuracy	Complexity and time-consuming to train and implement the model

Note. SVM = Support vector machine; GLM = Generalized linear model; Pls = Performance indicators; AFL = Australian Football League; NBA = National Basketball Association; XGBoost = eXtreme Gradient Boosting; LM = Left model; RM = Right model; HR = Heart rate; ML = Machine learning; VO₂ = Volume of oxygen; MO₂ = Muscle oxygen; ARNN = Autoregressive neural network

with their parameters, but also that the application involving cycling sport is the same as the one that will be conducted in the future. Furthermore, the data that they gave, which demonstrated a substantial correlation ($r = 0.94, p < 0.001$), further enhances our confidence in their empirical research.

DL METHOD OF TIME SERIES AND REAL-TIME PREDICTION

Artificial Neural Network (ANN)

ANN is a technique in AI that enables computers to interpret data in a manner analogous to the human brain (Amazon Web Services [AWS], n.d.). Ahmadi et al. (2024) demonstrated the application of this technique to estimate radial and mediolateral forces, offering a cost-effective solution for studying pedaling biomechanics using stationary cycle ergometers. This research revealed that a DL model can precisely forecast radial and mediolateral force components during cycling by utilizing cost-effective sensor data and participant attributes. The neural network model achieved an inter-subject normalized root mean square error (nRMSE) of 0.15 ± 0.02 and 0.26 ± 0.05 for radial and mediolateral forces, respectively, at high cadence, and 0.20 ± 0.04 and 0.22 ± 0.04 at self-selected cadence. The model exhibited diminished accuracy in forecasting mediolateral forces compared to radial forces, which could be enhanced by increasing the sample size and exploring a wider range of power outputs.

Additionally, Stetter et al. (2020), who utilized it to predict knee joint forces (KJF) during athletic activities, based on data acquired from wearable sensors. Thirteen individuals were outfitted with two inertial measurement units (IMUs) positioned on the right leg. Participants executed a range of actions, encompassing linear motions, directional changes, and jumps. Biomechanical modelling was conducted to ascertain KJF. Their model was developed to represent the correlation between the IMU signals and the KJF time series. The results demonstrated that the estimation accuracy of the ANN varied among movements, although it remained satisfactory for most of them.

Radial Basis Function Neural Network (RBFNN)

RBFNN, another variant of ANN, employs radial basis functions as activation functions. The RBFNN is a three-layered feedforward neural network. The initial layer is linear and solely disseminates the input signal, but the subsequent layer is nonlinear and employs Gaussian functions. The third layer linearly aggregates the Gaussian outputs. During training, only the weights connecting the hidden layer to the output layer are adjusted (Alireza, n.d.). Oytun et al. (2020) applied this method to predict specific types of athletic performance in female handball players, involving countermovement jumps with hands-free and hands-on-hips, 10-m and 20-m sprints, a 20-m shuttle run test, and a handball-specific

agility test. The study's results indicated that the radial-basis function neural network effectively predicted the examined forms of athletic performance, with R^2 scores ranging from 0.86 to 0.97.

Deep Neural Networks (DNN)

DNN is a type of ANN that has many deep layers, meaning there are many layers in between the input and output layers. Deep architecture facilitates the learning of more intricate models than shallow architecture. However, they also increase the quantity of significant training parameters, including the number of layers and the number of units per layer (Salvi et al., 2021). Zhao and Li (2023) proposed a study utilizing DNN to predict the risk of sports injuries in athletes following after-school physical exercise, which encompasses seven movements: squat, hurdle step, straight leg squat, shoulder flexibility, active straight leg lift, trunk stability push-up, and body rotation, involving 29 athletes. The validated DNN model achieved an accuracy of 88.10%.

J. Xu (2022) used the DNN method to analyze predictions of sports competition results, as accurate predictions enable educators to implement suitable training for students, hence optimizing learning outcomes. The model they tested has a significant impact on performance prediction, and the values generated by their prediction model can minimize systematic prediction errors, thereby enhancing the accuracy of performance forecasts for sports competitions in China, their primary focus. The extracted values, with an absolute prediction error of 1.6167 and an RMSE of 0.142, substantiated the efficacy of their AI model as claimed.

Zhou (2022) employed this strategy to forecast a player's performance accurately. Precise forecasting of athletic performance is beneficial not only for athletes but also aids in the advancement of sports. They assert that employing this approach has enhanced the accuracy of sports performance predictions and reduced prediction error in comparison to conventional methods. The prediction error value obtained with their AI model is 0.12, in contrast to the traditional method's value of 0.15.

Generative Adversarial Networks (GANs)

GANs are one of the types of ANN algorithms that are a category of neural networks that independently discern patterns within input data to produce novel instances that resemble the original dataset (GeeksforGeeks, 2025b). G. Li (2022) employed this method to assess the prediction of sports training performance based on physical fitness data of college students. The outcome achieved from their proposed model during the testing phase is 98% accuracy at the conclusion of the forecast. This demonstrated that their model attains superior categorization prediction outcomes compared to the other models they tested.

Convolutional Neural Network (CNN)

CNNs are a distinct kind of neural network specifically engineered to analyze grid-structured input, including images. They are especially adept at identifying and processing pictures. Inspired by the visual processing mechanisms of the human brain, CNNs proficiently capture hierarchical patterns and spatial dependencies in images (GeeksforGeeks, 2025a).

In 2024, G. Wang and Ren employed this strategy to precisely forecast college students' athletic performance and incentivize them to enhance their physical condition. The conclusion of their studies indicated an accuracy of 0.93, a precision of 0.94, and an F1-score of 0.93. Their test demonstrates the suggested model's superior predictive ability, validating its effectiveness in accurately forecasting sports achievements in accordance with actual criteria. From different perspectives, Pappalardo et al. (2020) offer a unique methodology for training a time series classifier utilizing a CNN to predict the likelihood of a player sustaining an injury in a future time frame, based on their historical workload in football. The results they achieved were 0.92 in precision and 0.88 in F1-score. This leads them to assert that their trials, conducted using injury records and workload data, validate the efficacy of the prediction model.

Imran (2022) suggested a CNN method for recognizing sporting activities to be utilized within wearable sensors. Wearable sensors can generate time-series data from accelerometers, gyroscopes, and magnetometers. Their AI model comprises 1,251 trainable parameters, achieving a test accuracy of 98.387% across six activities: badminton, basketball, cycling, football, skipping, and table tennis. The average F1 score, recall, and precision are 98.0, 98.7, and 98.0%, respectively.

LSTM

LSTM is an advanced variant of the recurrent neural network (RNN) that effectively captures long-term dependencies in sequential data, rendering them suitable for tasks such as language translation, speech recognition, and time series forecasting. In contrast to conventional RNNs that utilize a singular hidden state propagated through time, LSTMs incorporate a memory cell that retains information over prolonged durations, effectively tackling the issue of learning long-term dependencies (GeeksforGeeks, 2025e).

Nokihara et al. (2023) employed this methodology to analyze data, enabling players to gain an advantage in the rapid rally progression of badminton matches. They explored the innovative task of forecasting future shuttle trajectories in badminton match videos by detecting players and training a time-sequence model. They confirmed that the LSTM employed in this approach achieves maximum accuracy by transitioning to a time-sequence model. Two sorts of displacement errors were utilized as evaluation measures for the experiment. The first metric is the average displacement error (ADE), defined as the mean of the errors across all output frames. The second is the ultimate displacement error

(UDE). The mean value obtained for the ADE is 0.04908, whereas for the fractional-order differential equation (FODE), it is 0.08391.

Sadr et al. (2025) employed an LSTM technique to identify and predict sports injuries for the management of athletes' performance and health. The results obtained demonstrate that the LSTM model attained the best accuracy of 91.5%, confirming its efficacy in identifying and forecasting injury patterns in athletes. Their findings emphasize the potential for incorporating DL methodologies into sports injury management systems to improve early identification and preventive efforts.

The LSTM method was also employed by Hilmkil et al. (2018) to predict the heart rate response of a cyclist during a training session. Their LSTM model was trained for roughly one week on a Titan X Pascal GPU, achieving final minimum, mean, and maximum RMSE values of 2.51, 5.62, and 25.67, respectively, on the validation set. This demonstrated that it is feasible to train a model capable of predicting a cyclist's heart rate at any specific moment.

A different view of the idea taken by L. Wang et al. (2021) involves the integration of knowledge about big data-driven computer vision and DL algorithms, specifically LSTM, to facilitate intelligent labelling and representation of individual human movements depicted in video sequences for gymnastics. The results obtained from the LSTM model demonstrate a 14.22% improvement over the other models tested for sports video categorization and prediction.

Optimal athletic performance necessitates a balance between rigorous training and sufficient recovery. Because of this, Biró et al. (2024) employed LSTM to develop a resilient AI model aimed at forecasting weariness and stamina. Nineteen regular, injury-free runners from University College Dublin, Ireland, were used as a study dataset in the experiment, under both tiredness and non-fatigue conditions. The optimal outcome achieved from their evaluation of this model is a precision of 0.60, indicating that 60% of the predictions made were accurate. The recall and F1 score are both 0.60, reflecting precision and recall, which suggests a comparably balanced performance with a slight preference for accuracy. The results indicate that their LSTM model has reduced accuracy in forecasting fatigue levels, but they attribute this to observed performance improvements. The proposed enhancement methods may encompass additional hyperparameter optimization (modifying the quantity of LSTM units, altering the learning rate, or testing various activation functions), reassessing the feature engineering approach, or utilizing strategies to mitigate overfitting, which could be a concern due to the extensive number of training epochs in their research. The excessive number of epochs in the training process may lead to issues with computing efficiency and an increased potential for overfitting, especially if validation accuracy stabilizes early in the training phase. Consequently, it is essential to assess the model's performance on the validation set during the training phase to optimize the number of epochs and avert overfitting.

Others

Other researchers created their own models of AI or used hybrid methods or unspecified methods. One of the hybrid modifications of LSTM that is an outstanding idea comes from Zhang et al. (2022), which integrates the LSTM model with an attention mechanism and a time sliding window, resulting in the development of an attention-based sports-aware LSTM (AS-LSTM) designed to forecast outcomes in football matches. The results derived from their comparative table of actual and anticipated outcomes indicate that the model's accuracy reached 80% during the performance evaluation of the AS-LSTM model prediction. Consequently, they assert that their AS-LSTM model is accurate and helpful in predicting football match outcomes based on teams' past performance in competitions.

Watson et al. (2021) implemented an innovative hybrid concept model that utilized multi-input RNN and CNN to forecast the outcomes of sequences in Rugby Union, based on the ordered sequence of actions and their locations on the pitch. They assert that the CNN-RNN was the most effective single-input model across six outcomes when validated in real-world applications.

The soft actor-critic (SAC)-optimized DL model, developed by Tang et al. (2025), aims to enhance real-time monitoring and analysis of track and field athletes, overcoming the shortcomings of conventional monitoring systems regarding real-time performance and precision. The system was constructed based on a deep reinforcement learning algorithm. They have asserted that their study validates the considerable benefits of their technology over conventional approaches in terms of response speed, accuracy, and energy efficiency. Their proposed solution, the SAC-optimized system, enhances response time (200 ms compared to 250 ms) and data processing accuracy (98.5%) in real-time monitoring, especially for intricate sports such as high jump and discus throw.

Chen and Yuan (2021) applied a CNN algorithm hybrid with a residual network approach in their research on a sports injury prediction system utilizing visual analysis technology. They assert that their model has achieved one of the best accuracies on the training set, as evidenced by a final loss of only 0.6 at the conclusion of the forecast. This result demonstrates that their recommended strategy is effective.

Factorization deep product neural network (FDPNN) is a hybrid DL model aimed at enhancing predictive performance through the integration of factorization methods and DL (He & Chua, 2017). The primary components included are factorization machines (FM), product layer, and DNN (Guo et al., 2017). The researcher who incorporated this hybrid into their study was ZhaoriGetu (2022), to predict pupils' sports performance in school. They are evaluating various quantities of hidden layer neurons, with 256 emerging as the optimal choice, resulting in the FDPNN model achieving superior predictive performance in sports compared to alternative methods. The obtained values for accuracy, precision, and F1 score were 0.8668, 0.9499, and 0.9071, respectively.

R. Luo and Krishnamurthy (2023) introduces an innovative hybrid DL approach, termed GATv2-TCN, for forecasting player performance in basketball. This model is a hybrid AI framework that integrates a graph attention network with a temporal convolution layer, surpassing prior models that depend exclusively on player historical performance. The performance comparison reveals that the RMSE and MAE values for their model are 2.222 and 1.642, respectively, in contrast to other models examined in their study, which produced significantly higher values. They assert that this demonstrated its efficacy compared to other models they evaluated and is already yielding money in sports betting, while also offering useful information for sports analysis and predictions.

Precise sports forecasting is a crucial competency for professional coaches, enabling them to formulate effective training methodologies and strategic competitive approaches. Conventional approaches sometimes employ intricate mathematical statistical techniques to enhance predictability; nevertheless, these methods are often constrained by dataset size and struggle with long-term forecasts involving variable distributions, particularly poorly in forecasting point-set-game multi-level matches. To address this difficulty, H. Liu et al. (2024) propose TM^2 , a TCDformer-based Momentum Transfer Model for long-term sports prediction, which encompasses a momentum encoding module and a prediction module based on momentum transfer. Their analysis illustrates that TM^2 markedly surpasses current sports prediction models, achieving a 61.64% reduction in MSE and a 63.64% decrease in MAE, thus establishing a new standard in sports event prediction.

Another hybrid AI model idea was introduced by C. Liu et al. (2025), which integrated ARIMA and LSTM methodologies to conduct time series forecasting, specifically targeting the modeling of national medal counts to yield insights into prospective Olympic results. They achieved the final converged outcomes of RMSE and MAE values with consistent training cycles, specifically 0.098 for RMSE and 0.072 for MAE, which they assert are excellent results, demonstrating strong compatibility between ARIMA and LSTM in their hybrid AI model. Furthermore, their model may achieve an accuracy of up to 83% when utilizing a substantial volume of data. This demonstrated that their model is proficient and dependable for usage in this type of sports application.

The multidisciplinary characteristics of sports, along with numerous systemic and non-systemic components, complicate the prediction of match outcomes when relying on a singular disciplinary perspective. Consequently, Hsu (2021) conducted a study employing a DL methodology utilizing CNN and logistic regression for time series modelling to forecast sports match results. The primary data used in the prediction are sourced from betting market odds and real scores of each National Football League (NFL) game and are subsequently converted into sports candlesticks. The predictive efficacy of their model, CNN, was significantly enhanced by the proposed two-stage methodology, which utilized a logistic regression judgment model, achieving a prediction accuracy of 69% and surpassing that of the betting market.

Despite the valuable outcomes of data-driven algorithms in sports injury risk assessment, insufficient model generalization and the inability to automate feature extraction have hindered the practical application of study findings. Consequently, Ye et al. (2023) endeavor to develop an injury risk prediction model for running by integrating time-series picture encoding with DL techniques to tackle this issue more effectively. In comparison to the optimal model in their research, the proposed model shows enhancements in area under the curve (AUC), G_{mean} , sensitivity, and specificity by 23.9, 27.5, 39.7, and 16.2%, respectively, thus affirming the model's validity and applicability in injury risk prediction. The model presented in this paper provides a robust tool for predicting injury risk, aiming to enhance future sports injury prevention efforts.

A multitude of learning-based methodologies for efficient human activity recognition (HAR) have been recently established. Wearable inertial sensors are essential for HAR research to analyze sport-related activities. Sports activities are inherently unpredictable and have traditionally been categorized as complex, with standard ML methods employed to address HAR challenges. A hybrid DL model, designated as a multimodal bidirectional gated recurrent unit (MBiGRU) neural network, was introduced by Mekruksavanich and Jitpattanakul (2022) to identify commonplace sport-related actions, employing the publicly available Union Cycliste Internationale and Sports Dataset as a benchmark to evaluate the model's efficacy relative to other DL architectures in their research. The experimental findings suggested that the MBiGRU model achieved an accuracy rate of 99.55% and an F1-score of 99.55%, surpassing the baseline DL models in their study. Their proposed approach enhanced the efficacy of sensor-based HAR by utilizing the suggested MBiGRU paradigm.

Table 3 presents the best method among all those used in DL for predicting sports-related activities by implementing time series prediction, real-time prediction, or capturing temporal dependencies.

Summary

The entire journal above covered all research publications and journals that dealt with the application of DL techniques to the prediction of time series or real-time predictions for sports applications. The period covered by the paper examined spans from 2020 to 2025. Like in ML technique selection, a thorough examination of the application in sports, the performance of the AI models, and the outcomes that they achieved were the primary considerations that led to the selection of the findings from the comparison table and same as ML selection every one of the findings from the comparison table of DL will also serve as a guide for the subsequent study that will be conducted on this article, which is the construction of the Peephole-GRU Predictor for cyclist player performance prediction. A further observation that can be made is that between 2020 and 2025, a huge number

Table 3
Comparison of the best method by previous researchers based on the results obtained for implementing time series prediction and real-time prediction in deep learning methods

References	Algorithm method	Result		
		Model concept	Accuracy	Limitation
Hilmkil et al. (2018)	LSTM	Vanila	Final min, mean, and max RMSE of 2.51, 5.62, and 25.67 on the validation set	Time-consuming, which prevents them from exploring all modifications to the learning parameters
Zhang et al. (2022)	Attention-based sports competition performance prediction model (AS-LSTM)	Hybrid	The accuracy of the model reached 80% during its performance evaluation	A more appropriate sliding time window, a more effective word embedding method, and more accurate attention to athletes' psychological and physical states
Sadr et al. (2025)	LSTM	Vanila	Achieved the highest accuracy at 91.5%	The models require large and high-quality datasets for training, and incomplete or noisy data could significantly affect model performance
Tang et al. (2025)	SAC-optimized DL model	Hybrid	Data processing accuracy (98.5%) in real-time monitoring	Heavily reliant on high-performance hardware, which may limit its broader applicability in diverse environments, and the system may still encounter some errors and delays, indicating the need for further algorithm optimization in future research

Note. LSTM = Long short-term memory; RMSE = Root mean square error; AS-LSTM = attention-based sports-aware Long short-term memory; SAC = Soft actor-critic; DL = Deep learning

of researchers conducted series prediction or real-time prediction using the DL method. Furthermore, a huge number of researchers were also focused on a hybrid model of AI, rather than using the vanilla type. This may be caused by parameters such as accuracy and F1 score, which produce much higher values and a higher success rate in predicting sports elements than using a vanilla type of algorithm. The study by Zhang et al. (2022) became one of the main references (apart from the ML method) for our research. This is because their method of using the AS-LSTM model, which they have created and modified from the vanilla LSTM method, is one of the creative and innovative approaches to predicting sports competition results in football matches, thereby helping sports teams train and improve. This method was chosen from the best of the five models in the comparison table. The reason for this is not only that the model they have adapted is highly suited and compatible with their parameters, but also that the data they provide, which demonstrates the model's accuracy, reached 80% during their performance evaluation, further enhancing our confidence in their empirical research. Because this paper employs a hybrid method in the LSTM algorithm, it is also relevant to our future research, which will involve modifying an LSTM model into a GRU and a Peeppole model. Therefore, this paper is a suitable reference for our future research.

CONCLUSION

Sports analytics has been transformed by the inclusion of AI, especially ML and DL, into performance prediction, injury avoidance, and strategic decision-making. Particularly in cycling, basketball, and football, this paper presents a thorough study of time series and real-time prediction approaches used in various sports. Through an in-depth analysis of AI technologies from 2020 to 2025, both ML and DL models are crucial for enhancing sports analytics by providing more precise and data-driven insights. One important result of this study is the significant influence of AI on predictive modelling-based athlete performance optimization. High accuracy in forecasting outcomes and evaluating several performance criteria has been demonstrated by methods such as SVM, gradient boosting, LSTM networks, and hybrid models. Time series analysis enables athletes and coaches to make informed decisions based on real-time feedback, thereby guiding their choices of training and competition plans. AI applications, including VO_2 prediction, biomechanical analysis, and race strategy modelling, have given a competitive edge in cycling. The results show that hybrid AI architecture, LSTM-based models, and the LM + RM + HR approach all contribute to improved performance and reduced injury risk. Basketball and football have also seen similar developments, as AI-driven decision support systems enhance tactical planning, player scouting, and the prediction of game outcomes. There are still difficulties even with these developments. Using AI in sports requires large-scale databases, robust computing tools, and the ability to extend forecasts across multiple athletic

fields. Furthermore, the complexity of human performance introduces variability that AI models must constantly adjust to, emphasizing the necessity of further improvement and hybrid approaches that combine conventional sports science with AI-powered analytics. Furthermore, extending real-time AI applications from top-notch sports to grassroots levels may democratize performance statistics and provide amateur athletes and coaches with valuable insights. This demonstrates how quickly AI has become a transformative tool in sports analytics, as it provides unmatched performance prediction and clarity in strategic decision-making. The ongoing development of AI-driven techniques guarantees a data-driven approach to training, competition, and injury prevention, redefining how sports professionals evaluate and maximize performance. From all these studies, it is clear that every previous technique from past researchers will support and serve as a starting point for the development of the hybrid peephole-GRU predictor in this research, aiming to improve predictive capabilities in cycling sports.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support from the Fundamental Research Grant Scheme, under grant number FRGS/1/2024/ICT06/UNIMAP/02/1, from the Ministry of Higher Education Malaysia.

REFERENCES

- Aarons, M. F., Young, C. M., Bruce, L., & Dwyer, D. B. (2023). Real time prediction of match outcomes in Australian football. *Journal of Sports Sciences*, 41(11), 1115–1125. <https://doi.org/10.1080/02640414.2023.2259266>
- Ahmadi, R., Rasoulia, S., Veisari, S. F., Parsaei, A., Heidary, H., Herzog, W., & Komeili, A. (2024). A machine learning approach for predicting pedaling force profile in cycling. *Sensors*, 24(19), 6440. <https://doi.org/10.3390/s24196440>
- Aleza, M. E., & Vetrithangam, D. (2023). Use of artificial intelligence to avoid errors in referring a football match. In *International Conference on Artificial Intelligence and Applications Alliance Technology Conference* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICAIA57370.2023.10169463>
- Alireza. (n.d.). *Radial basis function neural networks (with parameter selection using K-means)*. MathWorks. <https://www.mathworks.com/matlabcentral/fileexchange/52580-radial-basis-function-neural-networks-with-parameter-selection-using-k->
- Almaleck, P., Massucco, S., Mosaico, G., Saviozzi, M., Serra, P., & Silvestro, F. (2024). Electrical consumption forecasting in sports venues: A proposed approach based on neural networks and ARIMAX Models. *Sustainable Cities and Society*, 100, 105019. <https://doi.org/10.1016/j.scs.2023.105019>
- Almarashi, A. M., 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 Science, Medicine and Rehabilitation*, 16, 28. <https://doi.org/10.1186/s13102-024-00815-7>

- Amazon Web Services. (n.d.). *What is a neural network?* AWS. <https://aws.amazon.com/what-is/neural-network/#:~:text=A%20neural%20network%20is%20a,that%20resembles%20the%20human%20brain>.
- Amendolara, A., Pfister, D., Settlemayer, M., Shah, M., Wu, V., Donnelly, S., Johnston, B., Peterson, R., Sant, D., Kriak, J., & Bills, K. (2023). An overview of machine learning applications in sports injury prediction. *Cureus: Journal of Medical Science*, 15(9), e46170. <https://doi.org/10.7759/cureus.46170>
- Anderson, J., & Rainie, L. (2023). *As AI spreads, experts predict the best and worst changes in digital life by 2035*. Pew Research Center. https://www.pewresearch.org/internet/2023/06/21/as-ai-spreads-experts-predict-the-best-and-worst-changes-in-digital-life-by-2035/?gad_source=1&gad_campaignid=22378837192&gbraid=0AAAAA-ddO9FyMKYfAtPXcL5wC6MOw4YCI&gclid=Cj0KCQjwrc7GBhCfARIsAHGeW5UC6B1lr3xK306xIWtAyon0_xnLrOiaQ8BCE8E7IQd6VokKSTzix28aAhtxEALw_wcB
- Atkinson, G., Davison, R., Jeukendrup, A., & Passfield, L. (2003). Science and cycling: Current knowledge and future directions for research. *Journal of Sports Sciences*, 21(9), 767–787. <https://doi.org/10.1080/0264041031000102097>
- Bangsbo, J. (2014). Physiological demands of football. *Sports Science*, 27(125), 1–6.
- Bangsbo, J. (2015). Performance in sports – With specific emphasis on the effect of intensified training. *Scandinavian Journal of Medicine and Science in Sports*, 25(S4), 88–99. <https://doi.org/10.1111/sms.12605>
- Biró, A., Cuesta-Vargas, A. I., & Szilágyi, L. (2024). AI-assisted fatigue and stamina control for performance sports on IMU-generated multivariate times series datasets. *Sensors*, 24(1), 132. <https://doi.org/10.3390/s24010132>
- Britannica. (2025). *Bayesian analysis: Statistics*. <https://www.britannica.com/science/Bayesian-analysis>
- Buchanan, R., Eliakim, R., & Eliakim, E. (2022). *Injury risk forecasting with Zone7 AI*. Zone7. <https://zone7.ai/case-studies/validation-study/validation-study-injury-risk-forecasting-with-zone7-ai/>
- Calleja-González, J., Mallo, J., Cos, F., Sampaio, J., Jones, M. T., Marqués-Jiménez, D., Mielgo-Ayuso, J., Freitas, T. T., Alcaraz, P. E., Vilamitjana, J., Ibañez, S. J., Cuzzolin, F., Terrados, N., Bird, S. P., Zubillaga, A., Huyghe, T., Jukic, I., Lorenzo, A., Loturco, I., ... Lago-Peñas, C. (2023). A commentary of factors related to player availability and its influence on performance in elite team sports. *Frontiers in Sports and Active Living*, 4, 1077934. <https://doi.org/10.3389/fspor.2022.1077934>
- Chen, X., & Yuan, G. (2021). Sports injury rehabilitation intervention algorithm based on visual analysis technology. *Mobile Information Systems*, 2021, 9993677. <https://doi.org/10.1155/2021/9993677>
- Cojocariu, I.-C. (2022). Predictive models applied in sports management – Literature review on research trends. *Journal of Public Administration, Finance and Law*, 11(23), 148–153. <https://doi.org/10.47743/jopaf-2022-23-11>
- Craig, N. P., & Norton, K. I. (2001). Characteristics of track cycling. *Sports Medicine*, 31, 457–468. <https://doi.org/10.2165/00007256-200131070-00001>
- Cui, J., Du, H., & Wu, X. (2023). Data analysis of physical recovery and injury prevention in sports teaching based on wearable devices. *Preventive Medicine*, 173, 107589. <https://doi.org/10.1016/j.ypmed.2023.107589>

- Dambroz, F., Clemente, F. M., & Teoldo, I. (2022). The effect of physical fatigue on the performance of soccer players: A systematic review. *PLOS One*, 17(7), e0270099. <https://doi.org/10.1371/journal.pone.0270099>
- De Beéck, T. O., Meert, W., Schütte, K., Vanwanseele, B., & Davis, J. (2018). Fatigue prediction in outdoor runners via machine learning and sensor fusion. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 606–615). Association for Computing Machinery. <https://doi.org/10.1145/3219819.3219864>
- Di Salvo, V., Adam, C., Barry, M., & Marco, C. (2006). Validation of Prozone®: A new video-based performance analysis system. *International Journal of Performance Analysis in Sport*, 6(1), 108–119. <https://doi.org/10.1080/24748668.2006.11868359>
- Dolci, F., Hart, N. H., Kilding, A. E., Chivers, P., Piggott, B., & Spiteri, T. (2020). Physical and energetic demand of soccer: A brief review. *Strength and Conditioning Journal*, 42(3), 70–77. <https://doi.org/10.1519/SSC.0000000000000533>
- Eid, A. I. A., Miled, A. B., Fatnassi, A., Nawaz, M. A., Mahmoud, A. F. A., Abdalla, F. A., Jabnoun, C., Dhibi, A., Allan, F. M., Elhossiny, M. A., Belhaj, S., & Mohamed, I. B. (2024). Sports prediction model through cloud computing and big data based on artificial intelligence method. *Journal of Intelligent Learning Systems and Applications*, 16(2), 53–79. <https://doi.org/10.4236/jilsa.2024.162005>
- Elmagd, A. M. (2019). General psychological factors affecting physical performance and sports. *Journal of Advances in Sports and Physical Education*, 2(7), 142–152. <https://doi.org/10.36348/JASPE.2019.v02i07.004>
- Faria, E. W., Parker, D. L., & Faria, I. E. (2005a). The science of cycling: Physiology and training – Part 1. *Sports Medicine*, 35, 285–312. <https://doi.org/10.2165/00007256-200535040-00002>
- Faria, E. W., Parker, D. L., & Faria, I. E. (2005b). The science of cycling: Factors affecting performance - Part 2. *Sports Medicine*, 35, 313–337. <https://doi.org/10.2165/00007256-200535040-00003>
- Fédération Internationale de Football Association. (2022). *Semi-automated offside technology to be used at FIFA World Cup 2022™*. FIFA. <https://inside.fifa.com/media-releases/semi-automated-offside-technology-to-be-used-at-fifa-world-cup-2022-tm>
- Fujii, K. (2021). Data-driven analysis for understanding team sports behaviors. *Journal of Robotics and Mechatronics*, 33(3), 505–514. <https://doi.org/10.20965/jrm.2021.p0505>
- Gavin, R. (2022). *Factors that effect sports performance: Sleep, mood, having a pre-match routine and romantic relationships* [Degree's thesis, National College of Ireland]. NORMA eResearch. <https://norma.ncirl.ie/5641/1/rorygavin.pdf>
- GeeksforGeeks. (2025a). *Convolutional Neural Network (CNN) in machine learning*. <https://www.geeksforgeeks.org/deep-learning/convolutional-neural-network-cnn-in-machine-learning/>
- GeeksforGeeks. (2025b). *Generative Adversarial Network (GAN)*. <https://www.geeksforgeeks.org/deep-learning/generative-adversarial-network-gan/> Retrieved (date) from url
- GeeksforGeeks. (2025c). *Gradient boosting in ML*. <https://www.geeksforgeeks.org/machine-learning/ml-gradient-boosting/>

- GeeksforGeeks. (2025d). *Support Vector Machine (SVM) algorithm*. <https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm/>
- GeeksforGeeks. (2025e). *What is LSTM - Long Short Term Memory?* <https://www.geeksforgeeks.org/deep-learning/deep-learning-introduction-to-long-short-term-memory/>
- Gomes, J., Portela, F., & Santos, M. F. (2013). Decision support system for predicting football game result. In X. Zhuang (Ed.), *Proceedings of the 19th International Conference on Computers, Recent Advance in Computer Science* (pp. 348-353). INASE Conferences. <https://www.inase.org/library/2015/zakynthos/COMPUTERS.pdf>
- Gribble, A. (2021). *Browns introduce Express Access at FirstEnergy Stadium*. Cleveland Browns. <https://www.clevelandbrowns.com/news/browns-introduce-express-access-at-firstenergy-stadium>
- Guo, H., Tang, R., Ye, Y., Li, Z., & He, X. (2017). *DeepFM: A factorization-machine based neural network for CTR prediction*. arXiv. <https://doi.org/10.48550/arXiv.1703.04247>
- Hastings, G. (2023). *IBM brings generative AI commentary and AI draw analysis to the Wimbledon digital experience*. IBM. <https://newsroom.ibm.com/2023-06-21-IBM-Brings-Generative-AI-Commentary-and-AI-Draw-Analysis-to-the-Wimbledon-Digital-Experience>
- He, X., & Chua, T.-S. (2017). Neural factorization machines for sparse predictive analytics. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 355–364). Association for Computing Machinery. <https://doi.org/10.1145/3077136.3080777>
- Hilmkil, A., Ivarsson, O., Johansson, M., Kuylensstierna, D., & van Erp, T. (2018). *Towards machine learning on data from professional cyclists*. arXiv. <https://doi.org/10.48550/arXiv.1808.00198>
- Hribernik, M., & Kos, A. (2023). Exploring the applicability of haptic actuators in aquatic environments. *Internet of Things*, 24, 100924. <https://doi.org/10.1016/j.iot.2023.100924>
- Hsu, Y.-C. (2021). Using convolutional neural network and candlestick representation to predict sports match outcomes. *Applied Sciences*, 11(14), 6594. <https://doi.org/10.3390/app11146594>
- Hu, X. (2024). Prediction of sports achievements based on time series analysis. In *International Conference on Cloud Computing, Performance Computing, and Deep Learning* (Vol. 13281, pp. 73–78). SPIE. <https://doi.org/10.1117/12.3050728>
- Ibrahim, M. R., Haworth, J., Christie, N., Cheng, T., & Hailes, S. (2017). Cycling near misses: A review of the current methods, challenges and the potential of an AI-embedded system. *Transport Reviews*, 41(3), 304–328. <https://doi.org/10.1080/01441647.2020.1840456>
- Imran, H. A. (2022). Khail-Net: A shallow Convolutional Neural Network for recognizing sports activities using wearable inertial sensors. *IEEE Sensors Letters*, 6(9), 1–4. <https://doi.org/10.1109/LESENS.2022.3197396>
- International Business Machines Corporation. (2021). *What is linear regression?* IBM. <https://www.ibm.com/think/topics/linear-regression>
- International Business Machines Corporation. (n.d.a). *What is random forest?* IBM. <https://www.ibm.com/think/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems.>

- International Business Machines Corporation. (n.d.b). *What is the k-nearest neighbors (KNN) algorithm?* IBM. [https://www.ibm.com/think/topics/knn#:~:text=The%20k%2Dnearest%20neighbor%20\(KNN\)%20algorithm%20is%20a%20non,used%20in%20macmach%20learning%20today](https://www.ibm.com/think/topics/knn#:~:text=The%20k%2Dnearest%20neighbor%20(KNN)%20algorithm%20is%20a%20non,used%20in%20macmach%20learning%20today).
- Karetnikov, A. D. (2019). *Application of data-driven analytics on sport data from a professional bicycle racing team* [Master's thesis, Eindhoven University of Technology]. TU/e Research Portal. <https://pure.tue.nl/ws/portalfiles/portal/139933495/ThesisAKaretnikov.pdf>
- Kolambe, M., & Arora, S. (2024). Forecasting the future: A comprehensive review of time series prediction techniques. *Journal of Electrical Systems*, 20(2s), 575–586. <https://doi.org/10.52783/jes.1478>
- Lastella, M., Roach, G. D., Halson, S. L., & Sargent, C. (2015). Sleep/wake behaviours of elite athletes from individual and team sports. *European Journal of Sport Science*, 15(2), 94–100. <https://doi.org/10.1080/17461391.2014.932016>
- Li, G. (2022). Construction of sports training performance prediction model based on a generative adversarial deep neural network algorithm. *Computational Intelligence and Neuroscience*, 2022, 1211238. <https://doi.org/10.1155/2022/1211238>
- Li, H., Cui, C., & Jiang, S. (2024). Strategy for improving the football teaching quality by AI and metaverse-empowered in mobile internet environment. *Wireless Networks*, 30, 4343–4352. <https://doi.org/10.1007/s11276-022-03000-1>
- Li, L., & Shi, X. (2014). Tennis “Hawk-Eye” technical research. *Journal of Chemical and Pharmaceutical Research*, 6(6), 298–305.
- Li, N., Hu, W., Ma, Y., & Xiang, H. (2024). Machine learning prediction of pulmonary oxygen uptake from muscle oxygen in cycling. *Journal of Sports Sciences*, 42(14), 1299–1307. <https://doi.org/10.1080/02640414.2024.2388996>
- Lingeswaran, S. (2023). *WSC Sports: How AI has led the highlights revolution*. Sportcal. <https://www.sportcal.com/features/wsc-sports-how-ai-has-led-the-highlights-revolution/>
- Liu, C., Ma, C., & Zhou, X. (2025). *Exploring patterns behind sports*. arXiv. <https://doi.org/10.48550/arXiv.2502.07491>
- Liu, H., Huang, X., Gu, J., Shi, J., He, N., & Feng, T. (2024). TCDformer-based momentum transfer model for long-term sports prediction. *Expert Systems with Applications*, 289, 128310. <https://doi.org/10.1016/j.eswa.2025.128310>
- Luo, R., & Krishnamurthy, V. (2023). *Who you play affects how you play: Predicting sports performance using graph attention networks with temporal convolution*. arXiv. <https://doi.org/10.48550/arXiv.2303.16741>
- Luo, S., Soh, K. G., Zhao, Y., Soh, K. L., Sun, H., Nasiruddin, N. J. M., Zhai, X., & Ma, L. (2023). Effect of core training on athletic and skill performance of basketball players: A systematic review. *PLOS One*, 18(6), e0287379. <https://doi.org/10.1371/journal.pone.0287379>
- Mekruksavanich, S., & Jitpattanakul, A. (2022). Sport-related activity recognition from wearable sensors using bidirectional GRU Network. *Intelligent Automation and Soft Computing*, 34(3), 1907–1925. <https://doi.org/10.32604/iase.2022.027233>

- Miah, J., Mamun, M., Rahman, M. M., Mahmud, M. I., Ahmad, S., & Nasir, M. H. (2022). MHfit: Mobile health data for predicting athletics fitness using machine learning models. In *2nd International Seminar on Machine Learning, Optimization, and Data Science* (pp. 584–589). IEEE. <https://doi.org/10.1109/ISMODE56940.2022.10180967>
- Mullin, B. (2016). *The Associated Press will use automated writing to cover the minor leagues*. Poynter. <https://www.poynter.org/tech-tools/2016/the-associated-press-will-use-automated-writing-to-cover-the-minor-leagues/>
- Nakisa, N., & Rahbardar, M. G., (2021). A review of the effect of psychological factors on soccer player's performance. *Indian Journal of Physical Education, Sports and Applied Science*, 11(1), 87–102.
- Noble, J. (2024). *What are ARIMA models?* IBM. <https://www.ibm.com/think/topics/arima-model>
- Nokihara, Y., Hori, R., Hachiuma, R., & Saito, H. (2023). Prediction of shuttle trajectory in badminton using player's position. In *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (Vol. 5, 788–795). SciTePress. <https://doi.org/10.5220/0011785800003417>
- NVIDIA Corporation. (n.d.). *XGBoost*. <https://www.nvidia.com/en-us/glossary/xgboost/>
- Oytun, M., Tinazci, C., Sekeroglu, B., Acikada, C., & Yavuz, H. U. (2020). Performance prediction and evaluation in female handball players using machine learning models. *IEEE Access*, 8, 116321–116335. <https://doi.org/10.1109/ACCESS.2020.3004182>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, 10, 89. <https://doi.org/10.1186/s13643-021-01626-4>
- Pantzalis, V. C., & Tjortjis, C. (2020). Sports analytics for football league table and player performance prediction. In *11th International Conference on Information, Intelligence, Systems and Applications* (pp. 1-8). IEEE. <https://doi.org/10.1109/IISA50023.2020.9284352>
- Pappalardo, L., Luca, G., Alessio, R., & Paolo, C. (2020). *Explainable injury forecasting in soccer via multivariate time series and Convolutional Neural Networks*. Barca Innovation Hub. https://static.capabiliaserver.com/frontend/clients/barca/wp_prod/wp-content/uploads/2020/01/c6658839-paper-format-luca-pappalardo-1.pdf
- Pennsylvania State University. (n.d.). *Introduction to GLMs*. <https://online.stat.psu.edu/stat504/lesson/6/6.1>
- Phillips, K. E., & Hopkins, W. G. (2020). Determinants of cycling performance: A review of the dimensions and features regulating performance in elite cycling competitions. *Sports Medicine - Open*, 6, 23. <https://doi.org/10.1186/s40798-020-00252-z>
- Plakias, S., & Michailidis, Y. (2024). Factors affecting the running performance of soccer teams in the Turkish Super League. *Sports*, 12(7), 196. <https://doi.org/10.3390/sports12070196>
- Proulx, D., & Smith, D. J. (2014). *Track skills development grid*. <https://cyclingbc.net/wp-content/uploads/2022/01/Cycling-Disciplines-Skills-Development-Grid.pdf>

- Ren, L., Wang, Y., & Li, K. (2024). Real-time sports injury monitoring system based on the deep learning algorithm. *BMC Medical Imaging*, 24, 122. <https://doi.org/10.1186/s12880-024-01304-6>
- Richman, P. (2018). *NBA 2013-14 Season: New SportVU system in arenas to improve player tracking*. Bleacher Report. <https://bleacherreport.com/articles/1764278-nba-to-install-sportvu-system-in-arenas-for-2013-14-and-improve-player-tracking>
- Ruddock, J. (2024). *Spiideo unveils industry-first product updates underscoring significant advancement in AI-automated sports broadcasting*. SVG Europe. <https://www.svg-europe.org/blog/news-roundup/spiideo-unveils-industry-first-product-updates-underscoring-significant-advancement-in-ai-automated-sports-broadcasting/>
- Sadr, M. M., Khani, M., & Tootkaleh, S. M. (2025). Predicting athletic injuries with deep learning: Evaluating CNNs and RNNs for enhanced performance and safety. *Biomedical Signal Processing and Control*, 105, 107692. <https://doi.org/10.1016/j.bspc.2025.107692>
- Salvi, M., Acharya, U. R., Molinari, F., & Meiburger, K. M. (2021). The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 128, 104129. <https://doi.org/10.1016/j.compbiomed.2020.104129>
- Sarker, I. H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2, 420. <https://doi.org/10.1007/s42979-021-00815-1>
- Seedhouse, E. (2017). Spaceports around the world. In *Spaceports around the world, a global growth industry* (pp. 101–113). Springer. https://doi.org/10.1007/978-3-319-46846-4_9
- Shukla, R., Jindal, P., Gupta, A., & Patil, H. Y. (2022). Total revenue prediction of a sports management application: Grook using machine learning models. In *13th International Conference on Computing Communication and Networking Technologies* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICCCNT54827.2022.9984472>
- Sim, J. (2024). *Global sports industry revenues to reach US\$260bn by 2033*. SportsPro. [https://www.sportspro.com/news/global-sports-industry-revenue-projection-2033-two-circles/#:~:text=Sports%20industry%20currently%20generates%20US\\$159%20billion%20in%20annual%20revenue,projected%20to%20outperform%20the%20market](https://www.sportspro.com/news/global-sports-industry-revenue-projection-2033-two-circles/#:~:text=Sports%20industry%20currently%20generates%20US$159%20billion%20in%20annual%20revenue,projected%20to%20outperform%20the%20market)
- Sports Business Journal. (2023). *Timeline of AI use in sports*. SBJ. <https://www.sportsbusinessjournal.com/Articles/2023/12/04/ai-timeline/>
- Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2020). Estimation of knee joint forces in sport movements using wearable sensors and machine learning. *Sensors*, 19(17), 3690. <https://doi.org/10.3390/s19173690>
- Stübinger, J., Mangold, B., & Knoll, J. (2020). Machine learning in football betting: Prediction of match results based on player characteristics. *Applied Sciences*, 10(1), 46. <https://doi.org/10.3390/app10010046>
- Suo, X., Tang, W., & Li, Z. (2024). Motion capture technology in sports scenarios: A survey. *Sensors*, 24(9), 2947. <https://doi.org/10.3390/s24092947>
- Taboga, M. (2021). *Bayesian estimation of the parameters of the normal distribution*. StatLect. <https://www.statlect.com/fundamentals-of-statistics/normal-distribution-Bayesian-estimation>.

- Tang, X., Long, B., & Zhou, L. (2025). Real-time monitoring and analysis of track and field athletes based on edge computing and deep reinforcement learning algorithm. *Alexandria Engineering Journal*, 114, 136–146. <https://doi.org/10.1016/j.aej.2024.11.024>
- Tingting, L. (2015). Research on the intelligent teaching system of college basketball based on artificial intelligence. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (18B), 49–60.
- Vec, V., Tomažič, S., Kos, A., & Umek, A. (2024). Trends in real-time artificial intelligence methods in sports: A systematic review. *Journal of Big Data*, 11, 148. <https://doi.org/10.1186/s40537-024-01026-0>
- Wang, F., & Zheng, G. (2022). What are the changes in basketball shooting pattern and accuracy in National Basketball Association in the past decade? *Frontiers in Psychology*, 13, 917980. <https://doi.org/10.3389/fpsyg.2022.917980>
- Wang, G., & Ren, T. (2024). Design of sports achievement prediction system based on U-net convolutional neural network in the context of machine learning. *Heliyon*, 10(10), e30055. <https://doi.org/10.1016/j.heliyon.2024.e30055>
- Wang, G., Pu, P., & Shen, T. (2020). An efficient gene bigdata analysis using machine learning algorithms. *Multimedia Tools and Applications*, 79, 9847–9870. <https://doi.org/10.1007/s11042-019-08358-7>
- Wang, L., Zhang, H., & Yuan, G. (2021). Big data and deep learning-based video classification model for sports. *Wireless Communications and Mobile Computing*, 2021, 1140611. <https://doi.org/10.1155/2021/1140611>
- Watson, N., Hendricks, S., Stewart, T., & Durbach, I. (2021). Integrating machine learning and decision support in tactical decision-making in rugby union. *Journal of the Operational Research Society*, 72(10), 2274–2285. <https://doi.org/10.1080/01605682.2020.1779624>
- WHOOP. (2023). *WHOOP unveils the new WHOOP coach powered by OpenAI*. <https://www.whoop.com/us/en/thelocker/whoop-unveils-the-new-whoop-coach-powered-by-openai/?srsltid=AfmBOorahCjNdPWXNgAA4VFS50SW73Z1XfwFg1Crzfc3tE37o43LzL22>
- Wohlwend, B. (2023). *Three regression models for data science: Linear regression, lasso regression, and ridge regression*. Medium. <https://medium.com/@brandon93.w/three-regression-models-for-data-science-linear-regression-lasso-regression-and-ridge-regression-6aac73c0d7a5>
- Wood, R. (2020). *The many sports of cycling*. Topend Sports. <https://www.topendsports.com/sport/cycling-sports.htm>
- Wright, C., Atkins, S., Jones, B., & Todd, J. (2013). The role of performance analysts within the coaching process: Performance analysts survey ‘The role of performance analysts in elite football club settings.’. *International Journal of Performance Analysis in Sport*, 13(1), 240–261. <https://doi.org/10.1080/24748668.2013.11868645>
- Wu, Y., Ma, Z., Zhao, H., Li, Y., & Sun, Y. (2020). Achieve personalized exercise intensity through an intelligent system and cycling equipment: A machine learning approach. *Applied Sciences*, 10(21), 7688. <https://doi.org/10.3390/app10217688>
- Xu, J. (2022). Prediction and planning of sports competition based on deep neural network. *Computational Intelligence and Neuroscience*, 2022, 1906580. <https://doi.org/10.1155/2022/1906580>

- Xu, T., & Tang, L. (2021). Adoption of machine learning algorithm-based intelligent basketball training robot in athlete injury prevention. *Frontiers in Neurorobotics*, 14, 620378. <https://doi.org/10.3389/fnbot.2020.620378>
- Yan, W., Jiang, X., & Liu, P. (2023). A review of basketball shooting analysis based on artificial intelligence. *IEEE Access*, 11, 87344–87365. <https://doi.org/10.1109/ACCESS.2023.3304631>
- Ye, X., Huang, Y., Bai, Z., & Wang, Y. (2023). A novel approach for sports injury risk prediction: Based on time-series image encoding and deep learning. *Frontiers in Physiology*, 14, 1174525. <https://doi.org/10.3389/fphys.2023.1174525>
- Yuan, X. (2024). Higher dimensional sports statistics and real-time game prediction. *Advances in Engineering Innovation*, 8, 9–18. <https://doi.org/10.54254/2977-3903/8/2024071>
- Zhang, Q., Zhang, X., Hu, H., Li, C., Lin, Y., & Ma, R. (2022). Sports match prediction model for training and exercise using attention-based LSTM network. *Digital Communications and Networks*, 8(4), 508–515. <https://doi.org/10.1016/j.dcan.2021.08.008>
- Zhao, J., & Li, G. (2023). A combined deep neural network and semi-supervised clustering method for sports injury risk prediction. *Alexandria Engineering Journal*, 80, 191–201. <https://doi.org/10.1016/j.aej.2023.08.048>
- ZhaoriGetu, H. (2022). Prediction of sports performance combined with deep learning model and analysis of influencing factors. *Scientific Programming*, 2022, 4082906. <https://doi.org/10.1155/2022/4082906>
- Zheng, D., & Yuan, Y. (2022). Time series data prediction and feature analysis of sports dance movements based on machine learning. *Computational Intelligence and Neuroscience*, 2022, 5611829. <https://doi.org/10.1155/2022/5611829>
- Zhou, Q. (2022). Sports achievement prediction and influencing factors analysis combined with deep learning model. *Scientific Programming*, 2022, 3547703. <https://doi.org/10.1155/2022/3547703>
- Zhu, Z. (2024). High dimensional sports statistics and machine learning in NBA. *Advances in Engineering Innovation*, 11, 78–94. <https://doi.org/10.54254/2977-3903/11/2024117>