International Journal on Robotics, Automation and Sciences

Review on Advancements in Artificial Intelligence and its Applications in Sports

Jun Jie Ooi*, Yit Hong Choo, Andi Prademon Yunus, Wei Hong Lim and Sui Yang Khoo

Abstract – The sport industry is being transformed by Artificial Intelligence (AI) in many ways. This paper seeks to discuss how AI has improved sports science, particularly in boosting the athletes' performance and avoiding injuries, through various machine learning models like Extreme Gradient Boosting, Support Vector Machines, and Random Forest Regression. These AI tools are more effective than the traditional methods, as they predict the athletes' performance results more accurately and managing their injuries more proactively. This paper also discusses the challenges of using AI in the sport industry, particularly in terms of data privacy and the reliability of the models. With the aid of AI, it is of no doubt that sport science will have a promising future.

Keywords—Athlete, Predict, Artificial Intelligence, Sports, Machine Learning.

I. Introduction

Artificial Intelligence (AI) is defined as a complex field involving the creation of computer systems that can learn from experience and adapt autonomously to produce predictive insights. As computational capabilities expand, the application and use of AI across various sectors, including sports has increased. The focus on assessing, mitigating, and preventing injuries is crucial, given their widespread occurrence and potential for severe physical, emotional, and financial impact, particularly at the professional level.

It is evident that AI has a significant impact across various industries, its integration has highlighted the limitations of traditional analytical techniques, which often rely heavily on the subjective judgements of coaches and basic statistical models that fail to capture the complex interactions of biomedical, physiological, and psychological data inherent in sport analytics [1]. To bridge this gap, researchers have increasingly turned to AI methodologies, such as ML models including Support Vector Machine and Random Forest Regression, which offer significant improvement in precision and reliability [2]. These methods can analyse large datasets with higher accuracy, providing insights into fatigue prediction, injury risk identification, and effective injury prevention strategies [3-7]. These advancements not only enhance predictive capabilities related to athlete performance but also significantly improve strategic decision-making in training and rehabilitation protocols, thereby expanding the boundaries of what is achievable in sports science and athlete management.

This study aims to explore deeper into the recent application of AI technologies in the field of sports, aiming to understand how AI is being utilised to enhance athletic performance, optimise training plans, and improve injury prevention strategies. The layout of this paper is organised as follows: Section 2 outlines

*Corresponding Author email: jjooi@deakin.edu.au ORCID: 0009-0000-5307-0350

Jun Jie Ooi* and Sui Yang Khoo are with the School of Engineering at Deakin University, Waurn Ponds, Australia (e-mail: jjooi@deakin.edu.au and sui.khoo@deakin.edu.au).

Yit Hong Choo is with the Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Waurn Ponds, Australia (e-mail: y.choo@deakin.edu.au).

Andi Prademon Yunus is with the Department of Informatics, Telkom University, Purwokerto, Indonesia (e-mail: andiay@telkomuniversity.ac.id).

Wei Hong Lim is with the Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur, Malaysia (e-mail: limwh@ucsiuniversity.edu.my).

International Journal on Robotics, Automation and Sciences (2025) 7, 1:58-63
https://doi.org/10.33093/ijoras.2025.7.1.7
Manuscript received: 24 Oct 2024 | Revised: 10 Dec 2024 | Accepted: 17 Dec 2024 | Published: 31 Mar 2025

Published by MMU PRESS. URL: http://journals.mmupress.com/ijoras

the methodologies used in this research; Section 3 provides the outcomes derived from the analysis; Section 4 discusses potential obstacles and issues that could arise with the further integration of AI in sports; Section 5 concludes with a synthesis of findings and the implications for the field of sports science. This structured approach ensures a comprehensive examination of the integration of AI into sports, facilitating a clearer understanding of its current capabilities and potential future development.

II. Methods

A. Search Strategy

A systematic electronic search was conducted to identify studies exploring the application of artificial intelligence methods in sports. The search covered all fields using the following terms: (“neural network” OR “machine learning” OR “deep learning” OR “artificial intelligence”) AND (“sports” OR “athlete” OR “players” OR “athletics”) AND (“performance” OR “fatigue” OR “injury” OR “monitoring” OR “prevention” OR “injury risk” OR “predict” OR “optimizing” OR “monitoring” OR “training load”).

B. Database

The literature search was conducted using the resources available on five electronic databases EBSCOhost, Academic Search Complete, MEDLINE, PubMed, and SPORTDiscuss.

C. Eligibility Criteria

After applying the search commands, the focus was narrowed to publications from 2018 onwards. This limitation was applied to ensure the research reviewed reflects the most current developments in the application of AI in sports. The criteria for inclusion were:

1. The study must be published in English.
2. The study must describe AI techniques or algorithms used.
3. Articles should quantitatively report results, showcasing measurable outcomes or findings.
4. Only include peer-reviewed articles to ensure the credibility and scholarly merit of the information.

D. Classifying the Main Research AI Technique or Method

When the review encounters research papers that employ multiple AI techniques, the technique demonstrating the best performance is prioritised. This approach focuses on the most effective methods within the study. This selection criterion ensures that the literature review highlights the most successful AI strategies, facilitating a deeper understanding of which techniques yield the best results in practical applications. This process simplifies extensive research findings into practical insights, highlighting the most effective and innovative AI techniques.

III. Results and Discussion

The review of the literature reveals significant advancements in the application of AI techniques to predict athlete performance and injury risk, highlighting the diverse methodologies and outcomes across various sports disciplines. The findings from several studies consistently demonstrate that AI models, particularly Support Vector Machines (SVM) and Random Forests (RF) [4, 8-10], have been effective in enhancing the predictability of athlete injuries and performance metrics when compared to traditional statistical methods.

In a study on predicting athlete fatigue from countermovement jump (CMJ) data, Wu, et al. [7] analysed force-time signatures collected using a force plate. Principal Components Analysis (PCA) and functional PCA (fPCA) were utilised to pre-process the data derived from ten recreational athletes performing CMJ at various intervals before and after training sessions. These techniques identified significant variations in force, time, and power metrics, particularly highlighting the first two principal components which accounted for 68% of the variations in the CMJ features. Linear Mixed Effects (LME) were subsequently used to predict neuromuscular fatigue, with the performance of models evaluated on their Mean Squared Error (MSE); the models achieved MSE values of 0.013 and 0.015 at 6- and 48-hours post-training, respectively.

Regarding injury prediction using AI, various approaches have been utilised, focusing on data collected through modern tracking technologies. Rossi, et al. [11] utilised global positioning system (GPS) training data capturing diverse physical activities and intensities. RF and Decision Tree (DT) models were chosen to analyse this data, with the DT model displaying a promising area under the curve (AUC) of approximately 0.76, indicating strong predictive capabilities.

Another investigation by Oliver, et al. [5] employed force plate to assess peak vertical ground reaction (VGRF) forces during various neuromuscular control test, including single-leg hop for distance, 75% hop and stick，and single-leg countermovement jump. This data was normalised to the body weights of athlete for analysis. The pre-processing of this data included discretization techniques to simplify continuous data into categories, enhancing the performance of various DT models such as Reduces Error Pruning Tree (REPTree), Alternating Decision Tree (ADT), and J48 consolidated (J48con). These models were assessed using metrics like AUC, with the J48con model achieving the highest performance with an AUC score of 0.663.

Moreover, Vallance, et al. [6] developed a tree-based algorithms, utilising non-linear models to predict injury risks of athlete based on a dataset collected including both external and internal load features. 10 Hz GPS system combined with a gyroscope, and a 100 Hz triaxial accelerometer were used to collect training data from soccer players, including external loads like number of accelerations and decelerations, total distance travelled, and maximum speed, as well as internal loads derived from perceived exertion rate and wellness assessments of athletes. The data underwent pre-processing which involved frequency (categorical variables) and mean (numerical variables) imputation, categorical variables were then converted into binary format. Extreme Gradient Boosting (XGB) appeared as the most effective models with an AUC of 0.93.

Mandorino, et al. [4] developed a ML approach to assess the neuromuscular status of player by predicting PlayerLoad (PL) from external load data. External load data， which was gathered through the WIMU Pro system. This system comprises a 3D magnetometer, three 3D gyroscopes, a 20 Hz ultra-wide band, and a 10 Hz GPS, allowing for comprehensive tracking of player metrics during both practices and games. The dataset includes various locomotor activities, for example distances travelled at various speeds and acceleration counts from 64 elite players monitored throughout the season. To eliminate potential factors that could mislead the results, partial sessions (where players trained below 90% of the entire session), individual sessions (where players trained separately), and rehabilitation sessions (as part of a recovery program) were excluded from the analysis. Additionally, categorical variables (position of players, and training phase) were transformed using one-hot coding method prior to utilising in the ML models. The performance of the RF model was notably high with a mean absolute percentage error (MAPE) of 0.10, indicating precise predictions of PL compared to actual measurements.

The research by White, et al. [10] explores the potential of using a AI model to predict peak external power in CMJ from data collected via a body-worn accelerometer. The ML model chosen was SVM, utilising VGRF data and accelerometer signals data for prediction, by achieving a peak power root mean squared error (RMSE) of 2.3 W/kg, which is about 5.1% of the mean, as determined through nested cross-validation and validated by an independent holdout test (2.0 W/kg).

In the study by Merrigan, et al. [9], eighty-two Division I NCAA football players were analysed to identify the most influential force-time metrics affecting CMJ height using RF. Data were collected using force plates that capture various force-time characteristics during CMJ tests, conducted without arm swings to reduce any interfering factors. The average absolute and relative metrics were included as predictor variables, with jump height designated as the dependent variable in the regression model. A similar overall variance in jump height was explained by the best RF models (8 metrics, R^2 = 0.95), despite using fewer metrics than stepwise regression (18 metrics, R^2 = 0.96). These findings may be used to inform training programs aimed at maximizing individual performance capabilities.

Gillett, et al. [12] assessed the effectiveness of a machine learning algorithm for identifying the key features that influence jump height. Bilateral ground reaction force (GRF) data were gathered from 89 right-handed male basketball athletes using force plates and associated software. An XGB model analysed fifty-six bilateral kinematic and kinetic variables from each condition to identify the top 10 most important features for predicting jump height and vertical jump reach (VJR) height using hands on hips (HOH) data. The model predicted VJR height from HOH jump data with a mean error of 7.13 cm.

This systematic review revealed that AI methods can be utilised to identify athletes at high risk of injury during sport participation and to identify risk factors. Although most of the studies reviewed using AI techniques to forecast injuries, their methodological quality varied from moderate to low. Given that sports performance is a growing field, there should be an encouragement for further developments in this promising area, given the substantial potential of AI techniques.

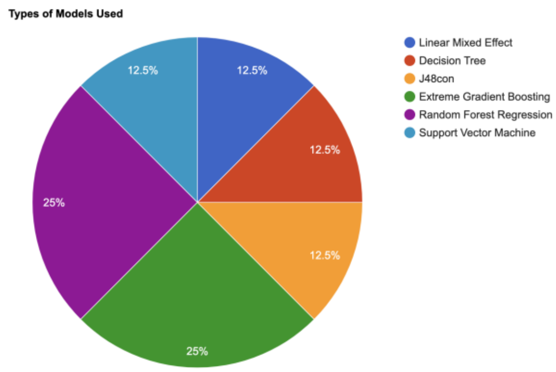

FIGURE 1. Types of AI models used.

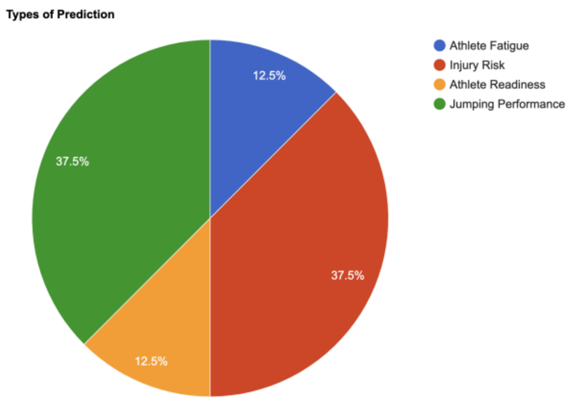

FIGURE 2. Types of prediction.

TABLE 1. Summary of AI used in sports.

Reference	Type of Prediction	Device	Data	Data Processing Method	N	Best Model	Metric	Performance
Wu, et al. [7]	Athlete fatigue	Force plate	CMJ were measured before, immediately, and at 0.5, 1, 3, 6, 24, and 48 hours after training.	PCA, fPCA	10	Linear Mixed Effects	MSE	0.013 at 6-hours and 0.015 at 48-hours
Rossi, et al. [11]	Injury risk	10 Hz GPS includes, 3D gyroscope, 3D digital compass, 100 Hz 3D accelerometer	Monitor players' physical activity for 23 weeks, placing devices between the scapulae through a tight vest.	12 features describing different aspects of the workload extract from GPS.	26	Decision Tree	AUC	0.76 ± 0.12
Oliver, et al. [5]	Injury risk	Force plate	Capture peak vertical ground reaction forces during neuromuscular test.	Discretization process.	355	J48con	AUC	0.663
Vallance, et al. [6]	Injury risk	10 Hz GPS includes gyroscope and 100 Hz triaxial accelerometer	External load (number of accelerations and decelerations, total distance travelled, and maximum speed) and internal load (perceived exertion rate and wellness assessment of athlete).	Mean imputation (numerical variables), frequency imputation (categorical variables), categorical variables converted into binary format.	40	Extreme Gradient Boosting	AUC	0.93
Mandorino, et al. [4]	Athlete readiness	WIMU Pro system (a 10 Hz GPS, a 3D magnetometer, three 3D gyroscopes, and a 20 Hz ultra-wide band)	External load data (distances covered at different speed, acceleration counts).	Only include training sessions that were fully completed, and categorical variables followed a one-hot coding protocol.	64	Random Forest Regression	MAPE	0.10 ± 0.01
White, et al. [10]	Jumping performance	Force plate (sampling frequency: 1000 Hz) and three Delsys Trigno sensors	Maximal effort CMJ with arm swings and without arm swings.	VGRF time series and accelerometer signals padded to the same duration.	69	Support Vector Machine	RMSE	2.3 W/kg
Merrigan, et al. [9]	Jumping performance	Force plate (sampling frequency: 1000 Hz)	2 maximal-effort CMJs without an arm swing.	Removed metrics that were less relevant in predicting CMJ height.	82	Random Forest Regression	R^2	93.1%
Gillett, et al. [12]	Jumping performance	Force plates (sampling frequency: 1000 Hz)	HOH jumps and VJR jumps.	Removed features related to landing, bilateral symmetry, and single-leg performance.	89	Extreme Gradient Boosting	R^2	91% during HOH jumps and 75% during VJR jumps

N: number of athletes; MSE: mean squared error; CMJ: countermovement jump; AUC: area under curves; MAPE: mean absolute percentage error; RMSE: root mean squared error

IV. Future Challenges

The adoption of AI in sports, while promising, faces substantial challenges that extend beyond the technical aspects. Many of the ML models in use today are unprotected to adversarial attacks. These attacks involve attackers intentionally manipulating the models to produce inaccurate predictions. In the study by Oseni, et al. [13], these attacks can occur in two phases, either training phase or testing phase of ML. Attackers have the ability to disrupt the learning process by injecting incorrect data into the training set during the training phase. During the testing phase, these attacks are known as evasion attacks. They use the weakness of model to create false inputs, tricking the model into incorrect predictions. This emphasises the need to integrate robust algorithms in the design and implementation of AI systems to shield against such threat.

Moreover, in this study mainly considered features derived from either pre-season lab tests (e.g., power and force) or data from wearable devices. However, employing wearables in sports analytics to measure both lower and upper body loads, along with inertial sensors placed on the upper body, remains the most common method for quantifying training loads. Various biomedical sensors were used to monitor the physiological load and biomechanical profiles of the athletes [14-16], although these sensors generally lacked AI capabilities. This could encourage greater investment in the research and development of sophisticated sensors, potentially that are fully integrated with AI capabilities, enhancing their effectiveness in the field of sports.

Financial constraints also play a crucial role, as the cost of developing and implementing sophisticated AI systems can be significant. This financial burden may limit the accessibility of AI technologies, particularly for underfunded sports organizations.

Those responsible for the design, execution, and implementation of AI-based applications in any field, including sports, must always consider regulatory guidance, liability, legal responsibility, and ethical considerations. Some standard and general ethics considerations applicable to AI include honesty, truthfulness, transparency, privacy, and safety. However, moral values and ethical standards are context-specific and, therefore, they might differ among populations, such as countries, regions, ethnics groups, etc.

V. Conclusion

This study demonstrates the transformative potential of AI within the sports industry. This extensive review has highlighted significant advancements in the application of AI to optimise athlete training, enhance performance, and mitigate injury risks. By developing sophisticated ML algorithms like RF, SVM, and XGB, researchers have significantly improved the predictability of performance outcomes and injury probabilities beyond traditional analytical techniques.

However, the integration of AI into sports is not without its challenges. Future efforts need to tackle privacy concerns, ensure ethical AI use in sports, adapt existing technologies, and enhance the reliability and interpretability of AI models. These steps will help maximise the deployment and utilisation of AI, allowing sports technologies to be implemented effectively and efficiently.

This study underscores the necessity for ongoing research and development in this field to fully leverage the capabilities of AI in enhancing sports science and athlete management. As AI technologies become more common and sophisticated, it is essential to develop comprehensive standards and protocols to ensure these tools are used ethically and effectively. The future of sports science hinges not only on technological advancements but also on the ability to integrate these innovations in a manner that respects the privacy and integrity of the data and the individuals it represents. The collaboration between technologists, researchers, and regulatory bodies will be essential to navigate the complexities of AI applications in sports, ensuring that the evolution of sports technologies contributes positively to the athletic and broader sports community.

Acknowledgment

We thank the anonymous reviewers for the careful review of our manuscript.

Funding Statement

There are no funding agencies supporting the research work.

Author contributions

Jun Jie Ooi: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation;

Yit Hong Choo: Project Administration, Writing – Review & Editing;

Andi Prademon Yunus: Project Administration, Supervision, Writing – Review & Editing;

Wei Hong Lim: Project Administration, Supervision, Writing – Review & Editing;

Sui Yang Khoo: Project Administration, Supervision, Writing – Review & Editing;

Conflict of interests

No conflict of interests were disclosed.

Ethics statements

Our research work follows The Committee of Publication Ethics (COPE) guideline. https://publicationethics.org.

References

[1] J. Borresen and M. I. Lambert, "The quantification of training load, the training response and the effect on performance," *Sports medicine,* vol. 39, pp. 779-795, 2009, DOI: https://doi.org/10.2165/11317780-000000000-00000

[2] S. Tedesco, S. Scheurer, K.N. Brown, L. Hennessy and B. O'Flynn, "A Survey on the Use of Artificial Intelligence for Injury Prediction in Sports," *2022 IEEE International Workshop on Sport, Technology and Research (STAR)*, pp. 127-131, 2022.
DOI: https://doi.org/10.1109/STAR53492.2022.9859939

[3] A. Cohan, J. Schuster, and J. Fernandez, "A deep learning approach to injury forecasting in NBA basketball," *Journal of Sports Analytics,* vol. 7, no. 4, pp. 277-289, 2021.
DOI: https://doi.org/10.3233/JSA-200529.

[4] M. Mandorino, A. Tessitore, C. Leduc, V. Persichetti, M. Morabito, and M. Lacome, "A New Approach to Quantify Soccer Players' Readiness through Machine Learning Techniques," *Applied Sciences,* vol. 13, no. 15, p. 8808, 2023
DOI: https://doi.org/10.3390/app13158808.

[5] J. L. Oliver, F. Ayala, M. B. D. S. Croix, R. S. Lloyd, G. D. Myer, and P. J. Read, "Using machine learning to improve our understanding of injury risk and prediction in elite male youth football players," *Journal of science and medicine in sport,* vol. 23, no. 11, pp. 1044-1048, 2020.
DOI: https://doi.org/10.1016/j.jsams.2020.04.021

[6] E. Vallance, N. Sutton-Charani, A. Imoussaten, J. Montmain, and S. Perrey, "Combining Internal- and External-Training-Loads to Predict Non-Contact Injuries in Soccer," *Applied Sciences,* vol. 10, no. 15, p. 5261, 2020.
DOI: https://doi.org/10.3390/app10155261

[7] P. P.-Y. Wu, N. Sterkenburg, K. Everett, D. W. Chapman, N. White, and K. Mengersen, "Predicting fatigue using countermovement jump force-time signatures: PCA can distinguish neuromuscular versus metabolic fatigue," *PLoS One,* vol. 14, no. 7, p. e0219295, 2019.
DOI: https://doi.org/10.1371/journal.pone.0219295

[8] I. M. K. Ho, A. Weldon, J. T. H. Yong, C. T. T. Lam, and J. Sampaio, "Using machine learning algorithms to pool data from meta-analysis for the prediction of countermovement jump improvement," *International Journal of Environmental Research and Public Health,* vol. 20, no. 10, p. 5881, 2023.
DOI: https://doi.org/10.3390/ijerph20105881

[9] J.J. Merrigan, J.D. Stone, J.P. Wagle, W.G. Hornsby, J. Ramadan, M. Joseph, S.M. Galster and J.A. Hagen, "Using Random Forest Regression to Determine Influential Force-Time Metrics for Countermovement Jump Height: A Technical Report," *Journal of Strength and Conditioning Research*, vol. 36, no. 1, pp. 277-283, 2022.
DOI: https://doi.org/10.1519/JSC.0000000000004154

[10] M. G. White, N. E. Bezodis, J. Neville, H. Summers, and P. Rees, "Determining jumping performance from a single body-worn accelerometer using machine learning," *Plos one,* vol. 17, no. 2, p. e0263846, 2022.
DOI: https://doi.org/10.1371/journal.pone.0263846

[11] A. Rossi, L. Pappalardo, P. Cintia, F. M. Iaia, J. Fernández, and D. Medina, "Effective injury forecasting in soccer with GPS training data and machine learning," *PloS one,* vol. 13, no. 7, p. e0201264, 2018.
DOI: https://doi.org/10.1371/journal.pone.0201264

[12] J. Gillett, J. De Witt, C. A. Stahl, D. Martinez, and J. J. Dawes, "Descriptive and kinetic analysis of two different vertical jump tests among youth and adolescent male basketball athletes using a supervised machine learning approach," *The Journal of Strength & Conditioning Research,* vol. 35, no. 10, pp. 2762-2768, 2021.
DOI: https://doi.org/10.1519/JSC.0000000000004100

[13] A. Oseni, N. Moustafa, H. Janicke, P. Liu, Z. Tari, and A. Vasilakos, "Security and privacy for artificial intelligence: Opportunities and challenges," *arXiv preprint arXiv:2102.04661,* 2021.
DOI: https://doi.org/10.48550/arXiv.2102.04661

[14] D.R. Seshadri, R.T. Li, J.E. Voos, J.R. Rowbottom, C.M. Alfes, C.A. Zorman and C.K. Drummond, "Wearable sensors for monitoring the physiological and biochemical profile of the athlete," *npj Digital Medicine*, vol. 2, no. 1, 2019.
DOI: https://doi.org/10.1038/s41746-019-0150-9

[15] R. De Fazio, V. M. Mastronardi, M. De Vittorio, and P. Visconti, "Wearable sensors and smart devices to monitor rehabilitation parameters and sports performance: an overview," *Sensors,* vol. 23, no. 4, p. 1856, 2023.
DOI: https://doi.org/10.3390/s23041856

[16] B. Sperlich, K. Aminian, P. Düking and H. Holmberg, "Editorial: Wearable Sensor Technology for Monitoring Training Load and Health in the Athletic Population," *Frontiers in Physiology*, vol. 10, 2020.
DOI: https://doi.org/10.3389/fphys.2019.01520