BALNEO RESEARCH Balneo and PRM Research Journal
English Edition
e ISSN 2734 - 8458 p ISSN 2734 - 844X ROMANIAN ASSOCIATION OF BALNEOLOGY

Research article

Athlete Injury Diagnosis System Using Machine Learning: A Case Study on Physiological and Athletic Performance Datasets

Firdausi Kahfi Maulana [1*], Ghulam Zaky Ismail [2], Susi Susanti [1], Tita Rachma Ayuningtyas [1], Isma Nur Azzizah [1] and Azizati Rochmania [2]

1 Department of Fisioterapi, Universitas Negeri Surabaya, Surabaya, Indonesia;
2 Department of physical education and health recreation, Universitas Negeri Surabaya, Surabaya, Indonesia

* Correspondence: firdausimaulana@unesa.ac.id

Citation: Maulana F.K., Ismail G.Z., Susanti S., Ayuningtyas T.R., Azzizah I.N., Rochmania A. - Athlete Injury Diagnosis System Using Machine Learning: A Case Study on Physiological and Athletic Performance Datasets
Balneo and PRM Research Journal ***2026**, 17(1): 998*

Academic Editor(s):
Constantin Munteanu

Reviewer Officer:
Viorela Bembea

Production Officer:
Camil Filimon

Received: 12.12.2025
Published: 31.03.2026

Reviewers:
Madalina Coman
Mariana Rotariu

Publisher's Note: Balneo and PRM Research Journal stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

BY-NC-ND

Copyright: © 2026 by the authors. Submitted for open-access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: Injuries among athletes pose a significant challenge in the world of sports, affecting not only physical performance but also long-term health and team dynamics. The ability to predict injury risk early is crucial in implementing preventive strategies. This study aims to develop predic-tive models for athletic injuries using machine learning algorithms, with a focus on comparing the performance of Logistic Regression and Random Forest on two datasets with differing class distributions, one balanced and one imbalanced. The first dataset (Injury Prediction Dataset) con-tains demographic and physiological data and has a balanced distribution between injured and non-injured classes. The second dataset (Collegiate Athlete Injury Dataset) contains performance and fatigue-related metrics, with a naturally imbalanced distribution where injury cases are rare. Both datasets underwent preprocessing, including feature scaling and label encoding, followed by model training and evaluation using accuracy, precision, recall, and F1-score metrics. Results show that Logistic Regression consistently outperforms Random Forest in detecting minority in-jury cases, especially in the imbalanced dataset, achieving 100% recall. In contrast, Random For-est fails to identify any injury cases despite high overall accuracy. This highlights the importance of model selection and handling class imbalance when building predictive systems for medical or safety-critical domains. The findings demonstrate that simple, well-tuned models like Lo-gistic Regression can be effective for early injury risk prediction, especially when supported by appropriate class balancing techniques. This research contributes to the development of AI-based medical diagnostic tools in the sports domain, offering a foundation for further enhancement us-ing richer datasets and advanced sampling methods.

Keywords: Athlete Injury Prediction, Machine Learning, Logistic Regression

1. Introduction

Injuries to athletes are among the most complex challenges in modern sports and physiotherapy. Injuries not only hinder physical performance and training continuity, but also impact psychological well-being, economic stability, and overall team success. [1], [2]. Epidemiological data show that musculoskeletal injuries account for more than half of absences in competitive sports, which have direct implications for rehabilitation processes, treatment costs, and the sustainability of athletes' careers. [3], [4]. Therefore, the ability to detect and predict injury risks early on has become a major focus in modern preventive sports medicine and physiotherapy.

Traditionally, injury risk identification has been conducted through subjective observation by coaches or physical therapists based on clinical experience and professional intuition. Although this approach has practical value, it relies heavily on the observer's perception and does not always have an objective quantitative basis[5],

[6]. In recent decades, advances in digital technology and the increased use of wearable sensor devices, GPS tracking systems, and physiological monitors have generated large volumes of data (big data) that can be systematically analyzed using artificial intelligence (AI) and machine learning (ML) approaches. The integration of these technologies opens up new opportunities to detect hidden patterns in data, connect various physiological indicators, and generate more accurate, adaptive, and large-scale injury risk predictions[7], [8].

Various studies have shown the success of machine learning in detecting health disorders such as cardiovascular diseases, neurological disorders, and musculoskeletal injury prognosis [9], [10], [11]. In the context of sports, Machine Learning algorithms have been used to model variables such as training load, fatigue levels, sleep quality, recovery duration, and injury history, enabling real-time monitoring of athlete conditions and providing adaptive feedback [12], [13]. However, the implementation of AI in sports injury prediction still faces several methodological challenges, such as data imbalance, limited physiological features, and low model interpretability, which implies a low level of clinical trust in the prediction results.

Most previous studies have focused on improving model accuracy, but few have examined in depth how data imbalance and model interpretability simultaneously affect prediction reliability. Complex models such as ensemble learning or deep learning often show high accuracy on balanced datasets but fail to detect rare injury cases in real-world data, where the proportion of injury cases is usually less than 10% of total training sessions [14], [15]. In addition, most of these models function as "black-box" systems that provide results without clear causal explanations, thus limiting their application in physiotherapy, which requires transparency and clinical reasoning in decision-making.

Seeing this research gap, an interpretable and explainable machine learning approach (Explainable AI/XAI) is required and should be able to adapt to various data distribution conditions. Two algorithms that stand out for this purpose are Logistic Regression (LR) and Random Forest (RF). Logistic Regression has the advantage in coefficient interpretation and clarity of relationships between variables, whereas Random Forest excels in capturing complex non-linear relationships. However, until now, there has been limited research systematically comparing the performance and stability of these two algorithms on balanced and imbalanced datasets in the context of athlete injury prediction. Based on these issues, this research has three main objectives, namely: 1) Developing an athlete injury prediction system based on artificial intelligence using physiological and performance datasets, 2) comparing the performance of Logistic Regression and Random Forest algorithms under balanced and imbalanced dataset conditions, 3) evaluating the contribution of key features such as exercise intensity, recovery time, and fatigue scores to improving the accuracy of prediction models.

By emphasizing model interpretability and performance reliability across diverse data distributions, this research is expected to bridge the gap between computational analysis and clinical application in sports physiotherapy. The results of this study are expected to contribute to the development of Explainable AI approaches in sports medicine and support the creation of transparent, accurate, and reliable preventive diagnostic tools for physiotherapists, sports scientists, and coaches in the early prevention of athlete injuries.

1. Materials and Methods

2.1 Research Design and Approach

This study uses a comparative experimental quantitative approach to develop and evaluate machine learning models in predicting athlete injury risk based on physiological and performance data. Two algorithms with different characteristics are used, namely Logistic Regression (LR) as an interpretable linear model, and Random Forest (RF) as an adaptive non-linear ensemble model capable of capturing complex relationships between variables. This approach is designed to assess the sensitivity of both models to two data distribution conditions, balanced and imbalanced, to evaluate the impact of data structure on injury detection capabilities.

2.2 Sources and Characteristics of Data

The research uses two open datasets from Kaggle.com that are relevant to the domain of exercise physiology and are open for replication.

1. Dataset 1 – Injury Prediction Dataset: Contains 1,000 athlete records with seven main features (Age, Height, Weight, Training Intensity, Recovery Time, Injury History, Likelihood_of_Injury). This dataset has a balanced target distribution between the "injured" (1) and "not injured" (0) classes, used to test model performance under ideal conditions.
2. Dataset 2 – Collegiate Athlete Injury and Performance Dataset: Consists of 200 athlete records with 17 physiological and performance variables such as Fatigue Score, Training Load, Heart Rate Variability, Recovery Index, and ACL Risk Score. Only about 7% of the data belong to the injury class, reflecting real-world conditions with high class imbalance.

Dataset selection was based on accessibility, physiological relevance, and suitability for replication research according to the FAIR Data principles (Findable, Accessible, Interoperable, Reusable).

2.3 Analysis Procedure

The analysis stages are conducted systematically through six main steps (Figure 1 – Research Workflow):

1. Data Exploration (EDA): Descriptive statistical analysis, outlier examination, and correlation heatmap visualization to understand data distribution and relationships between features.
2. Data Preprocessing: Includes label encoding for categorical variables, normalization with StandardScaler, and splitting data into training (80%) and testing (20%) sets using stratified sampling.
3. Model Training:
   a) *Logistic Regression:* Parameter class_weight='balanced' digunakan untuk mengoreksi proporsi kelas.
   b) *Random Forest:* Menggunakan 100 *estimators* dengan *max_depth* yang dioptimalkan melalui *grid search*.
2. Evaluasi Model:

Performa model diukur menggunakan *Accuracy, Precision, Recall, F1-score*, dan *Confusion Matrix*, dengan fokus utama pada *Recall* dan *F1-score* sebagai indikator sensitivitas terhadap kelas cedera.

3. Penanganan Ketidakseimbangan Data: Menggunakan pendekatan *class weighting* dan teknik *oversampling* (rencana SMOTE) untuk meningkatkan representasi kelas minoritas.

2.4 Interpretability Analysis

Using coefficient analysis (in LR) and feature importance (in RF) to assess the contribution of variables to injury prediction according to the principles of Explainable AI (XAI). As an alternative to SMOTE, class weighting and models sensitive to data distribution were directly employed, which still resulted in improved injury prediction performance.

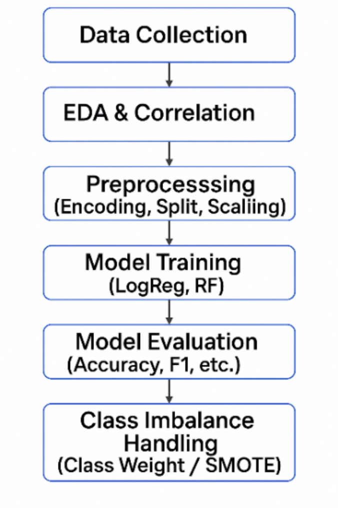

Figure 1 Research Workflow

2. Results

This study tested two machine learning models—Logistic Regression and Random Forest—with the aim of predicting athlete injuries based on physiological and performance data. Both models were trained and tested on two datasets with different characteristics: one balanced dataset (injury_data.csv) and one unbalanced dataset (collegiate_athlete_injury_dataset.csv).

Dataset 1: Injury Prediction Dataset (Balanced)

This dataset has 1000 data points with a balanced target distribution between injury (1) and no injury (0) classes. The model was trained with 800 data points and tested on 200 data points (80:20 split).

Table 1. Model Evaluation Results on Dataset 1 (Injury Prediction Dataset)

Model	Accuracy	Precision (C1)	Recall (C1)	F1-Score (C1)
Logistic Regression	52.5%	0.52	0.57	0.55
Random Forest	53.5%	0.53	0.57	0.55

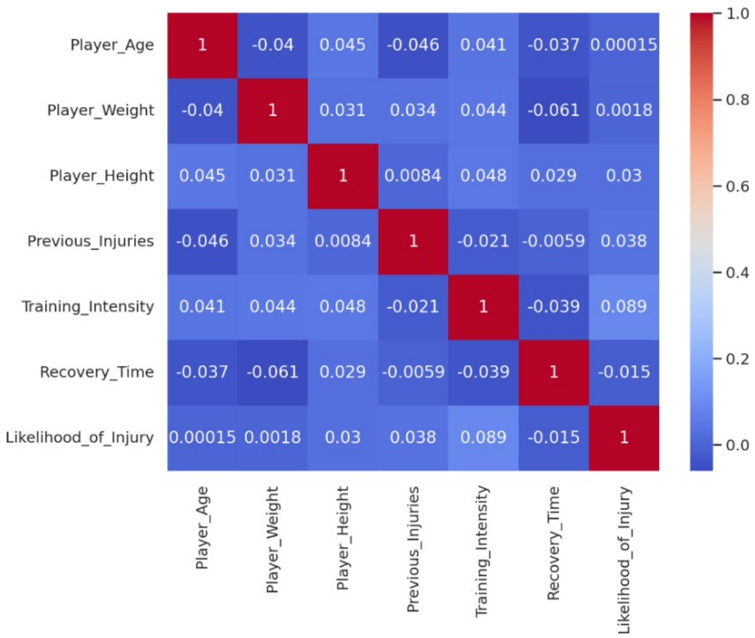

Figure 2. Dataset 1 Feature Correlation

In the first dataset (injury_data.csv), the distribution of targets between the injured and non-injured classes is relatively balanced (50% each). The two algorithms tested—Logistic Regression and Random Forest—show fairly comparable performance, with accuracy ranging from 52–53%. The F1-score for the injury class (1) also ranged around 0.55, which is relatively low. The interpretation of these results indicates that the features available in this dataset, such as Player_Age, Weight, Height, Training_Intensity, and Recovery_Time, do not yet have high predictive power to distinguish between injury and non-injury classes. The weak correlation between features and the target variable also supports this finding, as seen in the correlation heatmap. The linear model (Logistic Regression) and non-linear model (Random Forest) show nearly identical results, indicating that there may be no significant non-linear patterns in the data, or that the features are not sufficiently informative without further combination or transformation (e.g., feature engineering or feature interactions).

Note: The performance of both models is relatively low. Although the target classes are balanced, the input features may not be strong enough to produce sharp classifications. No significant differences were found between the linear model and the non-linear model (ensemble tree).

2. Dataset 2: Collegiate Athlete Injury Dataset (Unbalanced)

This dataset consists of 200 data points, with only 7% of the data labeled as injury (Injury_Indicator = 1). This poses a significant challenge because the model tends to be biased toward the majority.

The model was first trained with class weighting (adjusting the weight of the minority class), then retrained using a manual oversampling technique (similar to SMOTE).

Table 2 Model Evaluation Results on Dataset 2 (Collegiate Athlete Injury Dataset)

Model	Accuracy	Precision (C1)	Recall (C1)	F1-Score (C1)
Logistic Regression	95.0%	0.60	1.00	0.75
Random Forest	92.5%	0.00	0.00	0.00
LogReg + Oversampling	95.0%	0.60	1.00	0.75
Random Forest + Oversampling	92.5%	0.00	0.00	0.00

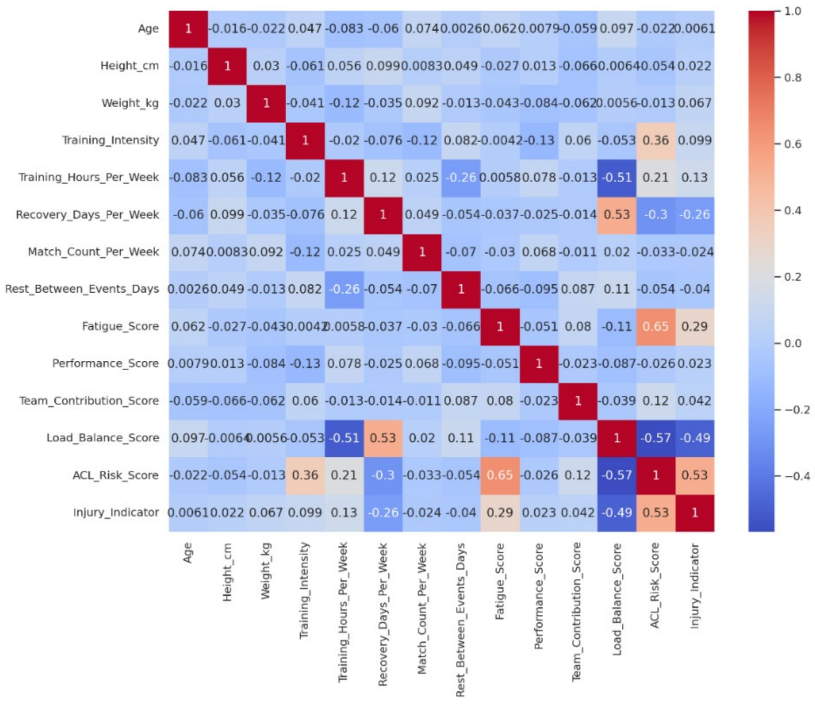

Figure 3 Feature Correlation Dataset 2

The second dataset (*collegiate_athlete_injury_dataset.csv*) presents a different challenge: a highly imbalanced target distribution, with only 3 out of 40 test samples (around 7%) belonging to the injury class. In such a scenario, accuracy is not a representative metric, as a model could appear "correct" simply by always predicting the majority class (no injury), while completely failing to identify injury cases.

In the initial experiment, Random Forest failed to identify any injury cases (recall = 0.0), even with adjusted class weights. This suggests that the model was heavily biased toward the majority class, a common characteristic of ensemble models when dealing with very limited minority class samples. In contrast, Logistic Regression performed better, achieving a recall of 100% and an F1-score of 0.75 for the injury class, despite a lower precision of 0.60. These results indicate that a linear model (with proper class weighting) is more sensitive and stable for small, imbalanced datasets. Oversampling techniques such as SMOTE, which was planned but not fully implemented, could be a viable next step to improve the minority class representation in the training data.

Notes:

1) Logistic Regression successfully identified all injury cases (recall = 100%), though with modest precision.
2) Random Forest tended to fail in classifying the minority class, even after class weighting.
3) This suggests that linear models are generally more stable for small, imbalanced datasets.

3. Comparison of the Two Datasets

The two datasets offer different insights:

1) Dataset 1 highlights the challenge of having less informative features, even though the target distribution is well-balanced.
   Dataset 2 emphasizes the issue of class imbalance, but features such as *Fatigue_Score*, *ACL_Risk_Score*, and *Training_Intensity* appear to be more relevant for detecting injuries.

4. Discussion

1. Model Performance Analysis

The research results show that the Logistic Regression (LR) model is capable of providing more consistent and sensitive performance compared to Random Forest (RF), especially in detecting injury cases in an imbalanced dataset. A recall value of 100% indicates that this model can identify all injury cases in the test data, although it is accompanied by an increase in false positives. In the context of physiotherapy and injury prevention, sensitivity (recall) is much more important than overall accuracy because failing to detect an injury (type II error) can have serious physiological consequences for athletes [16].

Conversely, Random Forest failed to detect the minority class in an imbalanced dataset despite achieving high accuracy (92.5%). This indicates a bias toward the majority class, a common problem in the application of ensemble algorithms when the number of injury data points is much smaller compared to normal data. This condition aligns with the findings of [17] who emphasized that class imbalance can hinder non-linear models in the context of training load-based injury prediction.

2. Physiological Relevance and Model Interpretability

The main advantage of the Logistic Regression model lies not only in its performance stability but also in its ability to be interpreted clinically. The coefficient analysis results show that Training Intensity and Fatigue Score have a positive influence on the increased risk of injury, while Recovery Time plays a protective role. These findings support the training–injury prevention paradox theory proposed by [18], [19], which explains that increasing training load without adequate recovery can lead to chronic fatigue, decreased tissue capacity, and ultimately increase the risk of musculoskeletal injuries.

The interpretability of this model becomes a crucial key in the context of Explainable Artificial Intelligence (XAI) in the field of physiotherapy. By knowing which factors most influence predictions, physiotherapists can carry out more targeted interventions, such as adjusting exercise intensity, extending recovery time, or enhancing fatigue monitoring strategies. This approach bridges the gap between computational intelligence and clinical practice, ensuring that AI-based decisions remain transparent and trustworthy. [20].

3. Practical Implications in Physiotherapy and Injury Prevention

From a practical perspective, interpretable injury prediction models have great potential to be integrated into AI-based athlete monitoring systems. These systems can provide early warnings of increased injury risk before clinical symptoms appear. For example, when the algorithm detects a combination of high Training Intensity and short Recovery Time, the system can alert coaches and physiotherapists to adjust the training load or conduct further assessments. This approach aligns with recent recommendations in sports physiotherapy, which advocate shifting the paradigm from treatment-based to prevention-based rehabilitation. With the help of such predictive models, physiotherapists can personalize interventions based on individual physiological responses, rather than merely following general training protocols. This enhances training efficiency while reducing the risk of recurrent injuries.

The findings of this study reinforce previous literature on the use of machine learning in sports. Studies by [21], [22], [23] utilized GPS data and training loads of professional soccer players to predict injury risk, with results showing that simple linear models often outperform complex models because they are easier to interpret and more stable against data variations. Logistic models with class weight adjustments provided more accurate results on small datasets with high data imbalance. Furthermore, this study introduces a new dimension by evaluating model performance in two different data distribution scenarios (balanced and imbalanced) as well as integrating the concept of explainable AI, which has not been widely applied in similar studies. This approach expands the understanding of how model stability and interpretability can be combined to produce physiologically and clinically relevant prediction systems. Although the results of this study show strong potential, there are several limitations that need to be considered. First, the dataset used is secondary and does not include biomechanical variables or real-time sensor data such as EMG, GPS, or plantar pressure, which could enrich the predictive model. Second, the amount of data is relatively small, especially for the injury class, so the generalization of the results is still limited. Third, this study has not applied synthetic oversampling techniques such as SMOTE or advanced interpretability methods such as SHAP and LIME, which could provide more granular explanations of the influence of individual features. Future research needs to integrate multimodal data (physiological, biomechanical, and clinical) and expand the sample size with athlete populations from various sports. In addition, cross-population external validation and field testing need to be conducted to assess the model's performance under actual training conditions. This approach will strengthen the application of AI that is transparent, ethical, and effective in supporting physiotherapists and coaches in preventing injuries in modern athletes.

5. Conclusions

This study successfully developed a machine learning-based athlete injury prediction model capable of classifying injury risk early through the analysis of physiological and performance indicators. The results show that prediction success highly depends on the quality and relevance of physiological features. A dataset that includes variables such as Fatigue Score, Training Intensity, and ACL Risk Score was proven to yield higher accuracy and sensitivity compared to a dataset containing only demographic data. It emphasizes that data-driven injury prediction is feasible and effective when supported by representative features. Comparison among algorithms shows that Logistic Regression consistently outperforms Random Forest, especially in detecting the minority class (injury) in imbalanced data. With a recall value reaching 100%, Logistic Regression proves to be more stable and efficient under limited data conditions, while Random Forest tends to be biased toward the majority class. In a balanced dataset, both models show relatively similar performance, indicating that model complexity does not always correlate with improved performance. These findings also confirm that data imbalance is a major challenge in predicting athlete injuries. The application of class weighting in Logistic Regression has proven to be a simple yet effective solution to reduce prediction bias. Further research is recommended to integrate the SMOTE (Synthetic Minority Oversampling Technique) method and adaptive feature selection to enhance the generalization and stability of the model on real-world data. Overall, simple interpretable models like Logistic Regression hold strategic value in physiotherapy and sports medicine, and open up opportunities for the application of Explainable AI (XAI) in accurate, transparent athlete monitoring systems that support preventive decision-making.

Author Contributions: Conceptualization, Firdausi Kahfi Maulana. and Ghulam Zaky Ismail.; methodology, Susi Susanti.; software, Tita Rachma Ayuningtyas.; validation, Isma Nur Azzizah., Azizati Rochmania. and Firdausi Kahfi Maulana formal analysis, Susi Susanti.; investigation, Ghulam Zaky Ismai; resources, Susi Susanti.; data curation, Tita Rachma Ayuningtyas.; writing—original draft preparation, Susi Susanti.; writing—review and editing, Firdausi Kahfi Maulana.; visualization, Ghulam Zaky Ismai.; supervision, Isma Nur Azzizah.; project administration, Firdausi Kahfi Maulana funding acquisition, Azizati Rochmania. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: This research article could not have been realized without the assistance of various parties. Therefore, the researcher would like to express the utmost gratitude to the entire research team from the Department of Physiotherapy and the Department of Physical Education and Health Recreation at Universitas Negeri Surabaya.

Conflicts of Interest: We acknowledge that there are no conflicts of interest related to this publication, and there has been no significant financial support for this work that could influence the results. As the corresponding author, I confirm that the manuscript has been read and approved for submission by all authors mentioned"

References

1. L. Ishøi, K. Krommes, R. S. Husted, C. B. Juhl, and K. Thorborg, "Diagnosis , prevention and treatment of common lower extremity muscle injuries in sport – grading the evidence : a statement paper commissioned by the Danish Society of Sports Physical Therapy (DSSF)," pp. 528–539, 2020, doi: 10.1136/bjsports-2019-101228.
2. S. Malaysiana, T. Sel, S. Mesenkima, K. Sukan, D. Kajian, and A. Klinikal, "Mesenchymal Stem Cell Therapy for Sports Injuries - From Research to Clinical Practice," vol. 49, no. 4, pp. 825–838, 2020.
3. J. M. Avedesian and W. Forbes, "Influence of Cognitive Performance on Musculoskeletal Injury Risk : A Systematic Review," 2021.
4. T. A. Buckley, C. M. Howard, J. R. Oldham, R. C. Lynall, C. B. Swanik, and N. Getchell, "No clinical predictors of postconcussion musculoskeletal injury in college athletes," Med. Sci. Sports Exerc., vol. 52, no. 6, p. 1256, 2020.

5. D. J. N. Wong et al., “Developing and validating subjective and objective risk-assessment measures for predicting mortality after major surgery: an international prospective cohort study,” PLoS Med., vol. 17, no. 10, p. e1003253, 2020.
6. P. Von Rosen and A. Heijne, “Subjective well-being is associated with injury risk in adolescent elite athletes,” Physiother. Theory Pract., vol. 37, no. 6, pp. 748–754, 2021, doi: 10.1080/09593985.2019.1641869.
7. M. Z. F. Khairuddin, K. Hasikin, N. A. Abd Razak, S. A. Mohshim, and S. S. Ibrahim, “Harnessing the multimodal data integration and deep learning for occupational injury severity prediction,” IEEE Access, vol. 11, pp. 85284–85302, 2023.
8. C. Rica, Y. Huang, S. Huang, Y. Wang, and Y. Li, “A novel lower extremity non-contact injury risk prediction model based on multimodal fusion and interpretable machine learning,” no. September, pp. 1–16, 2022, doi: 10.3389/fphys.2022.937546.
9. C. Munteanu, C. Popescu, V. Ciobanu, and G. Onose, “Biosignals and Motion-Derived Outcomes in Balneotherapy and Physical and Rehabilitation Medicine (PRM),” Balneo PRM Res. J., vol. 16, no. 3, pp. 1–15, 2025, doi: 10.12680/balneo.2025.885.
10. N. K. Iyortsuun, S. Kim, M. Jhon, and H. Yang, “A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis,” pp. 1–27, 2023.
11. R. Abd Rahman, K. Omar, S. A. M. Noah, M. S. N. M. Danuri, and M. A. Al-Garadi, “Application of machine learning methods in mental health detection: a systematic review,” Ieee Access, vol. 8, pp. 183952–183964, 2020.
12. F. Imbach, N. S. Charani, J. Montmain, R. Candau, and S. Perrey, “The Use of Fitness - Fatigue Models for Sport Performance Modelling : Conceptual Issues and Contributions from Machine - Learning,” Sport. Med. - Open, pp. 4–9, 2022, doi: 10.1186/s40798-022-00426-x.
13. F. Schwendinger et al., “Using machine learning–based algorithms to identify and quantify exercise limitations in clinical practice: are we there yet?,” Med. Sci. Sports Exerc., vol. 56, no. 2, p. 159, 2023.
14. H. Qi et al., “Automatic identification of causal factors from fall-related accident investigation reports using machine learning and ensemble learning approaches,” J. Manag. Eng., vol. 40, no. 1, p. 4023050, 2024.
15. A. Jamal et al., “Injury severity prediction of traffic crashes with ensemble machine learning techniques: A comparative study,” Int. J. Inj. Contr. Saf. Promot., vol. 28, no. 4, pp. 408–427, 2021.
16. L. D. Mendonça, C. Ley, J. Schuermans, E. Wezenbeek, and E. Witvrouw, “Physical Therapy in Sport How injury prevention programs are being structured and implemented worldwide : An international survey of sports physical therapists,” Phys. Ther. Sport, vol. 53, pp. 143–150, 2022, doi: 10.1016/j.ptsp.2021.06.002.
17. S. Bazarnovi and A. K. Mohammadian, “Addressing imbalanced data in predicting injury severity after traffic crashes: A comparative analysis of machine learning models,” Procedia Comput. Sci., vol. 238, pp. 24–31, 2024.
18. M. W. Dewangga, S. R. A. Dhani, R. A. Maghfiroh, E. Faozi, and S. I. Kurnia, “Cold Water Immersion’s Role in Enhancing Recovery Speed in Female Soccer Players After Matches: A Study of Physical Performance and Molecular Responses,” Balneo PRM Res. J., vol. 16, no. 2, pp. 1–11, 2025, doi: 10.12680/balneo.2025.821.
19. T. J. Gabbett, “Debunking the myths about training load, injury and performance: empirical evidence, hot topics and recommendations for practitioners,” Br. J. Sports Med., vol. 54, no. 1, pp. 58–66, 2020.
20. C. Dindorf, O. Ludwig, S. Simon, S. Becker, and M. Fröhlich, “Machine Learning and Explainable Artificial Intelligence Using Counterfactual Explanations for Evaluating Posture Parameters,” 2023.
21. 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,” Appl. Sci., vol. 10, no. 15, p. 5261, 2020.
22. C. Tiernan, T. Comyns, M. Lyons, A. M. Nevill, and G. Warrington, “The association between training load indices and injuries in elite soccer players,” J. Strength Cond. Res., vol. 36, no. 11, pp. 3143–3150, 2022.
23. G. Ravé, U. Granacher, D. Boullosa, and A. C. Hackney, “How to Use Global Positioning Systems (GPS) Data to Monitor Training Load in the ‘ Real World ’ of Elite Soccer,” vol. 11, no. August, 2020, doi: 10.3389/fphys.2020.00944.