S & M 4309

Integration of IoT and Sensor Technology in Sports Performance Tracking and Analysis

Jing Chen[1] and Wenbin Liu[2*]

[1]Shangqiu Normal University, Shangqiu 476000, China
[2]Fujian New World Computer Co. Ltd., Fuzhou 350105, China

(Received June 11, 2025; accepted January 15, 2026)

Keywords: IoT, sports performance, injury prediction, machine learning, wearable sensors, biomechanical analysis

We developed an IoT and machine learning (ML) system to predict injuries and monitor performance in athletes by integrating advanced sensor technology. IoT devices, including chest straps with heart rate monitors, inertial measurement units, accelerometers, and GPS trackers, were used to collect real-time physiological and biomechanical data. The data collected was analyzed using statistical methods and ML algorithms (Logistic Regression, Random Forest, and Extreme Gradient Boosting). The results showed that training load and fatigue are the most significant predictors of injury risk. While heart rate functioned as an independent marker, the participants under high strain showed significant cardiovascular overexertion with heart rate variability peaking between 100 and 200 BPM and median rates of 140 BPM. Ensemble ML models demonstrated exceptional predictive accuracy, reaching an area under the curve of 1.066. The results of this study demonstrate that the seamless integration of wearable sensors and data-driven analytics offers a robust approach to personalizing training and optimizing injury prevention.

1. Introduction

IoT has been widely used in sports. IoT is a network of devices and sensors that collect, process, and make decisions in real time. In sports, IoT technologies are applied in wearable devices, intelligent tools, and environmental sensors that continuously compile data on the athlete's movement and conditions.(1) By analyzing heart rate, movement speed, acceleration, and other physiological and physical data, coaches, trainers, and sports scientists gain important information about athletes' abilities and overall well-being. This enables them to fine-tune training strategies, optimize performance, and promote long-term health, which significantly benefits the sports community.(2)

Real-time data observation is a key to examining an athlete's overall performance. IoT technology ensures a seamless data flow to track an athlete's movement and biometrics during training and games. This allows for timely modifications to exercises or training methods,

*Corresponding author: e-mail: wsjy009@163.com
https://doi.org/10.18494/SAM5812

ISSN 0914-4935

helping prevent athlete overload and exhaustion. Real-time monitoring of athletes' biometric data and movements helps training programs to be tailored to an athlete's ability and performance, making training effective and efficient.[3] Environmental parameters such as temperature, humidity, and altitude are used to optimize performance and prevent injuries, which also enhances the athlete's performance. Diverse data used for athletes' healthcare are summarized in Fig. 1.

Machine learning (ML) algorithms have been widely used in sports medicine to predict and prevent injuries.[5] With IoT generating a vast amount of data, the algorithms identify complex patterns that are not discernible to humans. On the basis of data, ML models estimate the likelihood of injury, enabling preventive actions. This ability enables coaches to timely adjust training strategies according to the past performance of athletes, helps reduce the potential for sports injuries, and ensures fast recovery and resilience of athletes.

There are issues in incorporating IoT technology and ML for injury prediction and performance monitoring: effectively and seamlessly integrating the technologies, ensuring the accuracy of data collected through advanced sensor technology, and constructing ML algorithms. Therefore, we developed a system for the seamless integration of IoT and ML to predict injuries and track performance in this study.

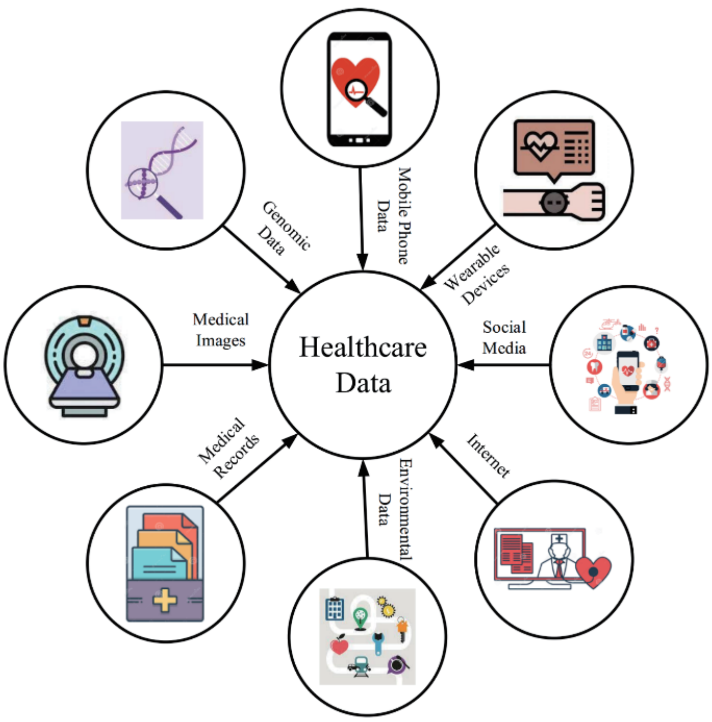

Fig. 1. (Color online) Diverse data sources used in sports.[4]

2. Theoretical Background

2.1 IoT on performance tracking and athlete engagement

IoT technology enables accurate reporting on athletes' physiology and movements, which helps trainers and athletes make immediate adjustments. Wearable devices, such as smartwatches, fitness bands, and sensor-embedded clothing, are used to monitor heart rate, oxygen saturation rate, speed, and movement.(6) Coaches need to adjust training programs to improve performance and reduce injuries in athletes.

Catapult and WHOOP© are high-tech GPS tracking systems and biometric wristbands. They are used to study past and current results and make exercise plans tailored to each athlete. In sports clothes, accessories, balls, and rackets, IoT technology is embedded to track the impact of training and competitions on athletes. In modern sports, enhancing performance with continuous data analysis is considerably emphasized. Embedded IoT technology also enables post-injury care. Wearable sensors collect data related to early signs of fatigue, insufficient hydration, or unsafe movement patterns that can lead to injury.(7)

IoT technology helps improve safety in sports, as coaches and athletes can prevent injuries and remain involved in sports for longer periods by using the data collected. For example, IoT devices monitor muscle activity and the type of movement, enabling an efficient recovery process after injury (Fig. 2). By monitoring athletes' biometrics, coaches adjust training and competition strategies.

IoT technology also increases athlete engagement by facilitating communication and enabling personalized coaching. It helps athletes and coaches exchange information in real time, regularly monitor performance, and set goals based on monitored progress in training and competition.(9)

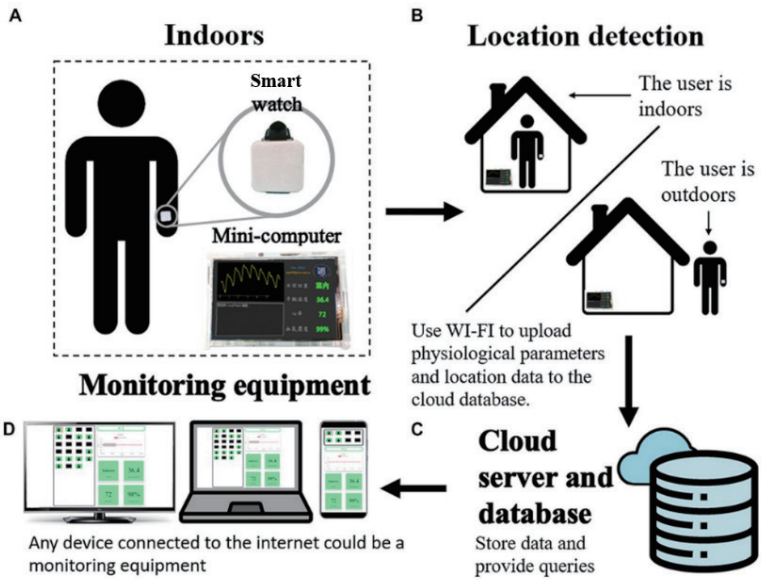

Fig. 2. (Color online) IoT-based wearable health monitoring device and its operations.(8)

Through IoT analytics, coaches track each athlete's performance to provide appropriate support and encourage athletes to be more responsible for teamwork. Moreover, IoT technology helps adjust strategies and tactics based on data updated in real time. Coaches can make informed decisions about tactics based on the ongoing actions of athletes and the team. As IoT analytics also help identify what opponents struggle with, coaches and athletes can make informed decisions to win. IoT technology enables reliable and ethical training and competition.

Sensors play an important role as they are essential in monitoring athletes' performance and physiological status. The results are important to make tactical adjustments during a match.[10] In addition, sensors implemented in sports equipment and venues are used to accurately measure distance, speed, and time without errors. For example, sensors in the starting block accurately detect a false start.

IoT and sensor technologies have led to technological advancements in sports, providing reliable measurement, information, and decision,[11] which are essential for fair play. Referees are supported by technology for reliable and fair decisions. Such technologies also enable fans to participate actively and enhance their experience and interaction. With sensors and network-connected devices, spectators can instantly access real-time statistics and detailed information about teams and athletes. In various applications, sports events are broadcast live, allowing viewers to enjoy key moments and exciting features repeatedly. These features make sports fans more engaged and devoted to sports games. By leveraging data analytics and IoT technology, broadcasters and event organizers can enhance their content and advertising strategies, delivering a more engaging experience for viewers while maximizing revenue opportunities. The integration of advanced technology in sports benefits all stakeholders, from athletes and teams to fans and sponsors.

Nevertheless, limitations exist in the widespread adoption of IoT technology in sports. Data privacy and security are concerns to be addressed. False data and disinformation must be screened as they have negative effects. High costs to implement the technology are also a hurdle for small teams and amateur athletes. Resistance to the introduction of advanced technology has also been observed. Therefore, effective data governance, regular platform updates, and education on the benefits of IoT technology adoption must be provided to address these issues for the more effective use of IoT technology in sports.

IoT technology enhances team performance and game fairness based on accurate data. Injury prediction and detection, personalized training, team strategies, and the fairness of competitions can be enhanced by using wearables, smart devices, and advanced data, which leads to better training and competition results. Spectators and viewers can better enjoy sports. However, data privacy, financial constraints, and resistance to the adoption need to be addressed for the more widespread use of IoT technology in sports.

2.2 Biomechanical factors in injury prediction

To determine the likelihood, extent, and type of injuries in sports, understanding the biomechanics of athletes is mandatory. Biomechanics encompasses the study of movement, force, and body responses to movements, which is a basis for improving safety or decreasing the risk of injury. Analyzing movement patterns in sports helps identify potential risks and prevent injuries, ensuring that athletes perform safely and efficiently. Performing repetitive routine exercises can strain muscles over time, while errors in executing precise movements increase the risk of injuries. For example, incorrect running form, such as overextending the front foot or excessive bouncing, causes stress fractures, plantar fasciitis, or other running-related problems. Repeated tossing or hitting can disrupt scapulohumeral rhythm, especially when shoulder motion or strength is imbalanced, increasing the risk of tendinopathy and shoulder impingement.(12) A weak core and limbs can cause excessive pressure on the lower back when twisting, increasing the risk of lumbar spine injuries.

Muscle repetitive stress injuries often arise from overstretching, excessive tension during muscle contraction, and imbalances between muscle groups (Fig. 3).(13) Since eccentric movements subject muscles to intense strain and structural changes, injuries frequently occur under eccentric loading conditions. Muscle damage and inflammation are affected by the degree of muscle over-lengthening during maximum stretch or rapid movements. The strength and coordination disparities of muscle groups might cause injuries as they compromise joint stability and movement efficiency. For example, weak hip stabilizers combined with overly strong quadriceps can increase knee valgus and the chances of anterior cruciate ligament injuries during cutting motions. Using biomechanical models and advanced motion tracking systems, clinicians assess joint flexibility, analyze muscle elongation, and monitor joint-loading patterns to enhance injury prevention and rehabilitation.

Overuse injuries in running-related sports often result from variations in ground reaction forces, stride mechanics, and joint movement patterns. The force and speed at which the foot strikes the ground significantly affect injury potential, as rapid landings increase stress on bones and joints. Excessive pronation or supination during the stance alters weight distribution, elevating the likelihood of stress fractures, plantar fasciitis, and Achilles tendinopathy.(15) Additionally, excessive internal hip movements, common in sports such as soccer and basketball, increase the risk of noncontact lower-body injuries. These movement inefficiencies often stem from neuromuscular fatigue or improper mechanics, which are detected through detailed motion analysis. Optimizing cadence, step length, and muscle strength substantially reduces injury

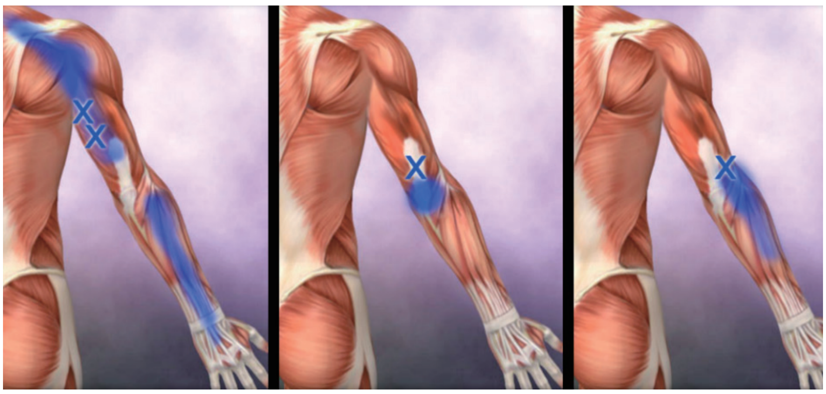

Fig. 3. (Color online) Repetitive strain injury.(14)

rates. Therefore, these biomechanical causes must be understood for effective injury prevention and rehabilitation.

Recent advancements in biomechanical analysis with ML have enhanced injury prediction and the identification of key neuromuscular and biomechanical markers. A study assessing elite young soccer players pinpointed concentric knee extensor peak torque, hip moments during single-leg drop jumps, and center-of-pressure sway as primary predictors of noncontact lower-extremity injuries. Despite a moderate accuracy [the area under the receiver operating characteristic curve (AUC) = 0.63], the study showed the importance of neuromuscular control and biomechanical stability in assessing injury risk using the least absolute shrinkage and selection operator regression.(16,17) Similarly, ensemble models, such as Extreme Gradient Boosting (XGBoost), demonstrated a 78% precision in predicting hamstring strain injuries by analyzing hamstring muscle strength and stiffness.(18) These models integrated biomechanical sensor data with physiological indicators, such as heart rate variability and respiratory patterns, as important information. These studies presented the importance of using advanced analytics and biomechanical data in improving injury risk assessment and prevention strategies.

Biomechanical modeling enables personalized injury prevention plans by simulating training programs. For example, targeted exercises that focus on the eccentric (downward) phase of movement can enhance muscle resilience against high-impact forces and reduce injury susceptibility. Strengthening opposing muscle pairs and incorporating dynamic warm-ups can improve joint stability and overall mobility. Reducing high-intensity movements and extreme ranges of motion minimizes microtrauma and supports proper tissue recovery. These data-driven interventions not only successfully lower injury rates but also enhance athletic performance. By integrating IoT-connected real-time biomechanics with advanced ML analytics, injuries can be predicted and prevented effectively.

3. Materials and Methods

We integrated IoT technology and ML algorithms into the monitoring of athletes' biometric data and movements. Wearable IoT devices were used to collect data to identify the patterns and relationships of the data collected. The performance and injury factors of the athlete were evaluated through such data collection and advanced data analytics to ensure robust and reproducible results to develop evidence-based training and injury prevention methods.

This study was conducted in three phases: (1) a literature review and pilot testing to select appropriate IoT devices and algorithms with optimal accuracy and reliability, (2) the system integration of sensor calibration, data source identification, and ML algorithm selection, and (3) system performance and usability evaluation (Fig. 4). The research phases collectively ensured the robustness and applicability of the developed system. In this study, we designed and implemented an IoT system architecture that encompassed data capture, storage, processing, and visualization. The system underwent validation with a cohort of 200 track and field athletes, encompassing a range of movements such as sprinting, jumping, and change-of-direction (COD) maneuvers.(19)

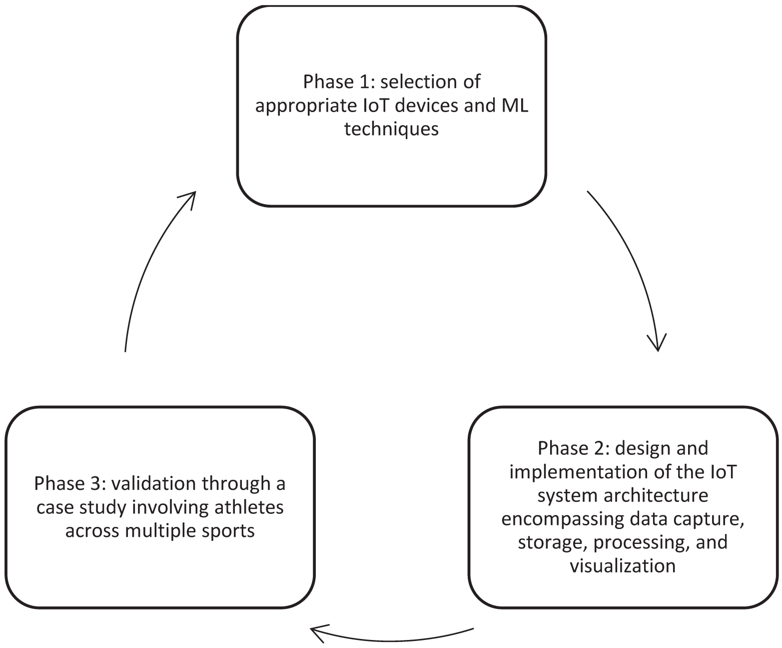

Fig. 4. Research process in this study.

3.1 Data collection

Data were collected using a chest strap in which a heart rate monitor, an inertial measurement unit (IMU), an accelerometer, a gyroscope, and a GPS tracker were embedded. The devices were selected, considering their measurement accuracy, data transmission speed and stability, and compatibility with commercial software.(19) Heart rate variability was tracked as it is an important indicator of stress and recovery. IMU and an accelerometer were used to assess injury risks based on biomechanical data, such as joint angles, body speed, and types of movement. Temperature and humidity were used to assess the effects of environmental factors on performance changes. The collected data were transmitted via Bluetooth Low Energy to a cloud server, which used a nonrelational database architecture. The data were captured during training and competitions for the monitoring of athletes' health conditions and performance in various events.

3.2 Data analysis

The data collected were analyzed using statistical methods and ML clustering algorithms. The noise and artifacts of the raw sensor data were removed using a Butterworth low-pass filter, defined by the following transfer function *H*.

$$H(S) = \frac{w_c^2}{s^{2+\sqrt{w_c s}+w_c^2}} \tag{1}$$

Here, ω_c is the cutoff frequency, and s is the complex frequency variable. Equation (1) was used to ensure the integrity of movement data for subsequent analysis.

The participants were athletes and exposed to a high risk of injury, so a uniform risk profile was expected. Therefore, unsupervised learning was prioritized to identify subtle biomechanical deviations and physiological markers of overexertion. K-means and hierarchical clustering algorithms were used to classify movement variability patterns for the detection of abnormal movement signatures, such as compensatory joint angles or altered ground reaction forces. These are early indicators of potential injury without the necessity of labeled injury outcomes. To predict these risk-associated biomechanical states, supervised ML models (Logistic Regression, Random Forest, and XGBoost) were trained to identify high-strain states characterized by excessive training loads and physiological fatigue. The predictive capability was evaluated using AUC and F1-scores to ensure the robustness of the system in identifying these precursor states across different athletic subgroups.

In this study, SPSS was used for statistical data analysis, and Python programming for ML model development and validation. SPSS was used to obtain descriptive statistics and conduct correlation and regression analyses to explore relationships between physiological variables and injury risks. Python libraries, such as NumPy, Pandas, scikit-learn, and TensorFlow, were employed to implement supervised learning algorithms (logistic regression, random forest, and XGBoost) for injury prediction.[20] Logistic regression was used to estimate the probability of injury occurrence (P) as

$$P = \frac{1}{1 + e^{-(B_0 + B_1 X_1 + B_2 X_2 + B_0 \cdots + B_n X_n)}}, \tag{2}$$

where B_0 is the intercept, B_1 is a coefficient, and X_n is a predictor of training load and biomechanical metrics. Random forest and XGBoost algorithms were used for ensemble-based classification as they have superior accuracy and robustness against overfitting.

We calculated the training load in arbitrary unit (arb. unit) using the session-rating of perceived exertion (sRPE) method. This involved multiplying the athlete's RPE (on a modified Borg CR-10 scale) by the total duration of the session in minutes. This dimensionless unit enables the standardized comparison of workload across various training modalities and is a validated metric for monitoring fatigue and injury risk in professional sports.

$$\textit{Training Load}\ (\mathrm{AU}) = \textit{Intensity}\ (\mathrm{RPE\ Scale\ 1–10}) \times \textit{Duration}\ (\mathrm{min}) \tag{3}$$

Here, intensity is measured using the Borg CR-10 scale (where 1 is very easy and 10 is maximal effort), and duration is the total time of the training session in minutes.

GPS tracker data was utilized to quantify external training load by monitoring high-frequency movement variables, including total distance covered, sprint distance, and movement velocity. These metrics were integrated with accelerometer-derived mechanical load and physiological heart rate data to create a comprehensive profile of an athlete's physical exertion. Specifically, velocity and distance data from the GPS sensors served as critical inputs for the

ML models, allowing the system to correlate high-intensity movement patterns with the onset of physical fatigue and subsequent injury risk.

3.3 Ethical considerations

This study was conducted following the ethical guidelines established by the institutional review board of Shangqiu Normal University before recruiting participants and collecting data. The participants were notified of the study goals, steps, possible harms, and advantages, and submitted their written permission. Personal information was excluded from the whole data, and stored and protected by regulations. Each participant was allowed to quit the experiment whenever they wanted. The use of IoT devices was monitored to prevent any discomfort or issues that might affect athletes' performance. It was ensured that ML models were used to process the collected data legally, and decisions on using data were made openly and fairly. This study was conducted while obeying the highest ethical standards in sports science research.

4. Results and Discussion

4.1 Factors affecting injury risk

The correlation matrix (Fig. 5) presents the quantitative relationships among physiological and biomechanical factors used to assess athlete health and highlights the importance of an integrated monitoring approach. Injury risk is not attributable to a single isolated factor but emerges from the interaction of multiple variables.

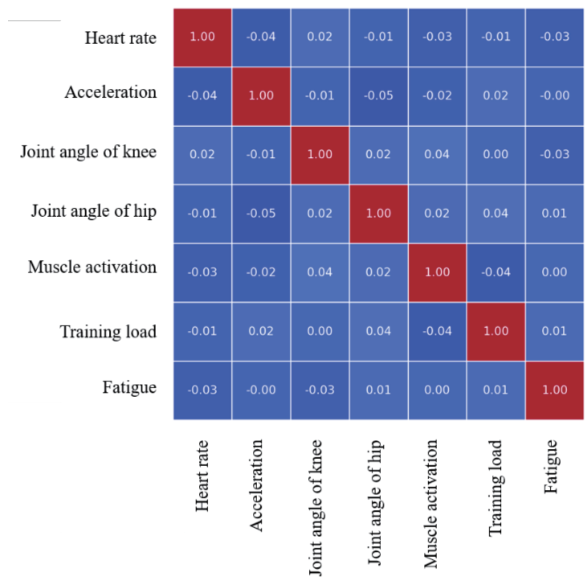

Fig. 5. (Color online) Correlation matrix of factors affecting injury risk.

Heart rate exhibits negligible correlation with other assessed variables ($r < 0.042$), indicating that it functions as an independent physiological marker rather than directly reflecting specific biomechanical movements. Nevertheless, heart rate variability is a critical indicator of cardiovascular stress. The participants experiencing high physiological strain displayed elevated heart rate variability (100–200 BPM) compared with the baseline data (120–160 BPM), with outliers exceeding 180 BPM identified as states of cardiovascular overexertion, resulting from insufficient recovery. Figure 6 shows the heart rate distribution of the participants. The box in the plot represents the interquartile range (IQR). The bottom edge of the box corresponds to the first quartile ($Q1$), or the 25th percentile, which is approximately 130 BPM. The top edge of the box corresponds to the third quartile ($Q3$), or the 75th percentile, which is approximately 150 BPM. The total height of the box represents the IQR, encompassing the middle 50% of the data. The horizontal line inside the box indicates the median, or the 50th percentile. In this case, the median heart rate is exactly 140 BPM. The vertical lines, commonly referred to as whiskers, extend from the box to the highest and lowest values that are not considered outliers. Typically, these whiskers extend to values within $1.5 \times IQR$ from the quartiles. Figure 6 shows that the median heart rate is approximately 140 BPM, with a majority of observations falling between 130 and 150 BPM.

A moderate positive correlation ($r = 0.15$) was observed between training load and fatigue, reflecting the expected biological response to exercise intensity, whereby increased training loads contribute to the accumulation of physical fatigue. This relationship confirms that higher exercise intensity is a primary driver of accumulated physiological strain. Knee and hip joint angles demonstrate limited interaction with other lower extremity joints. Moreover, the absence of significant correlations between muscle activation and joint angles suggests that neuromuscular control remains relatively stable and is not strongly dictated by specific joint positions in this study.

Training load was closely associated with injury risk. The participants exhibited training loads between 60 and 125 units, whereas their baseline training loads ranged from 15 to 80 units

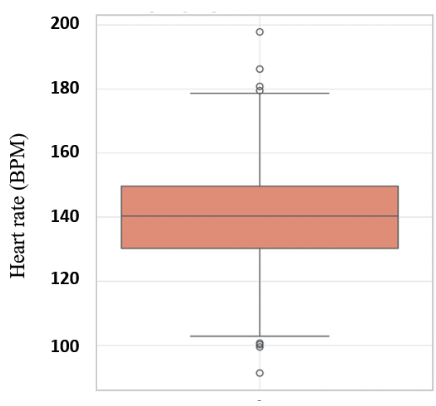

Fig. 6. (Color online) Distribution of heart rate (BPM) measurements.

(Fig. 7). Specifically, the high-risk group trained with a load of approximately 80 units, which was 60% greater than the median of 50 units observed in the low-risk group. Extreme training with limited recovery substantially increases the likelihood of injury, and loads exceeding 120 units indicate overreaching or overtraining, potentially impairing tissue adaptation and recovery. These findings emphasize the necessity of integrating individualized training load monitoring into injury prevention programs to safeguard athlete health and performance.

The generally weak correlations among factors underscore the appropriateness of the selected feature set for ML algorithms. High correlations (multicollinearity) can obscure predictive clarity, whereas the distinct nature of these variables enables algorithms to accurately weigh the unique contribution of each factor in injury prediction. The results suggest that heightened injury vulnerability arises from the integration of high training loads and physiological instability rather than from a single determinant.

4.2 ML model performance

ML models showed varying levels of accuracy in injury prediction, while ensembled models outperform traditional logistic regression in terms of accuracy. In this study, the prediction accuracy of Random Forest and XGBoost models was 1.0, exceeding logistic regression with an accuracy of 0.99 (Fig. 8). Ensemble models identified complex relationships of multiple factors influencing injury risk. Their high accuracy suggested that they are well-suited for analyzing the intricate interactions affecting injury prediction. However, the potential for overfitting must be considered, particularly due to the limited number of observations. Logistic Regression, Random Forest, and XGBoost models showed high performance, with their receiver operating characteristic (*ROC*) close to 1.0 for identifying differences in injury risk between high- and low-injury-risk groups. The models predicted well, although their sensitivities and specificities varied. The *ROC*s of logistic regression, random forest, and XGBoost models were 1.0, 1.0, and 0.99, respectively (Fig. 9). The XGBoost model's *ROC* (0.99) was slightly lower than that of the

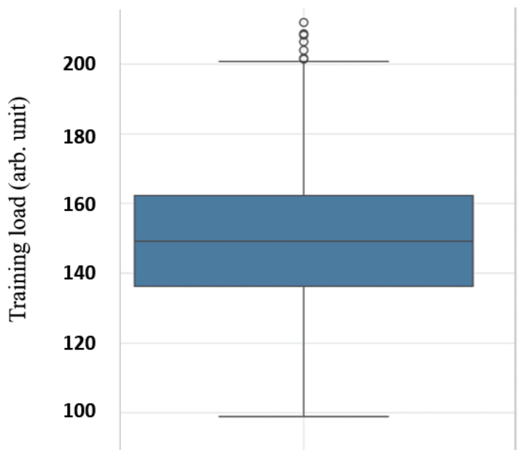

Fig. 7. (Color online) Distribution of training load (units) measurements.

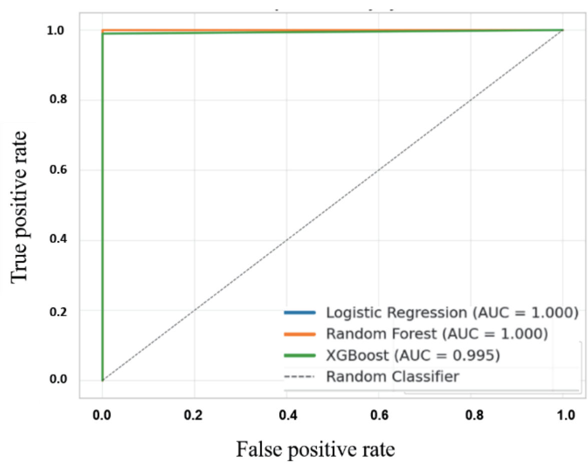

Fig. 8. (Color online) Receiver operating characteristics of injury prediction by ML models in this study.

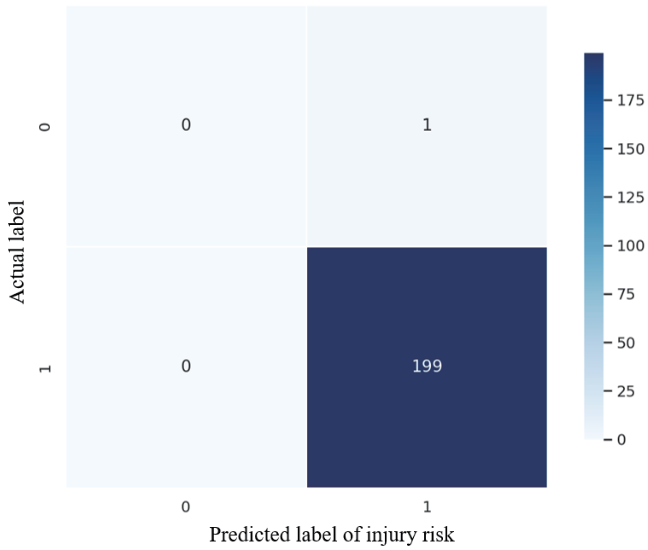

Fig. 9. (Color online) Confusion matrix of logistic regression model.

random forest model, primarily owing to its regularization techniques, which limit performance in certain cases.

The confusion matrix of the logistic regression model is presented in Fig. 9. The confusion matrix reveals that the predictive model is exceptionally accurate, but the single different result provides an important insight into its limitations. For one participant, the model predicted a low risk even though the participant belonged to the high-risk group. Statistically, this single error reduces the Recall from a perfect 1.0 to 0.995. The model correctly identified 199 participants as being at high risk, which are recorded as True Positives. This high number indicates that the model is extremely effective at recognizing the patterns associated with injury risk in the majority of the population. The accuracy of the model is 99.5%, and the recall (or sensitivity) is also 99.5%. Because the model did not incorrectly flag anyone as high risk who was actually not, the precision is 100%. The only false negative is a significant result, as the participant who needed intervention or rest was classified as safe.

4.3 Feature importance

Feature importance was calculated to identify factors playing an important role in injury risk prediction. Feature importance was attributed to training load and fatigue. In random forest and XGBoost models, the feature importance of training load was the largest, accounting for approximately 0.6 of the entire predictive power. The intensity and frequency of exercise were important in predicting an injury (Fig. 10). A feature importance score of fatigue of the XGBoost model was 0.38, indicating a close relationship between fatigue and injury.

A moderate feature importance of knee and hip joint angles was observed, indicating that they were not primary predictors, although they played an important role in injury prediction. The joint angle of the knee was more important than that of the hip, as knee injuries are more common in athletes. Muscle activation contributed minimally to injury prediction, implying that broader movement mechanics and loads on muscles were stronger indicators of injury risk than neuromuscular parameters. While heart rate regulation is important, it was not a significant predictor owing to its variability and indirect relationship with injury risk. Monitoring training loads is more effective in predicting injury than physiological metrics.

Random forest and XGBoost models accurately predicted injury, relying on similar features. The XGBoost model emphasizes complex cases owing to its gradient-boosting process, which continuously identifies and refines weak prediction abilities. Both models highlighted training load as the most influential factor in injury prediction. By integrating more physiological data, an injury prediction method can be enhanced. However, the results of this study provide a reference for the development of IoT monitoring systems across various sports.

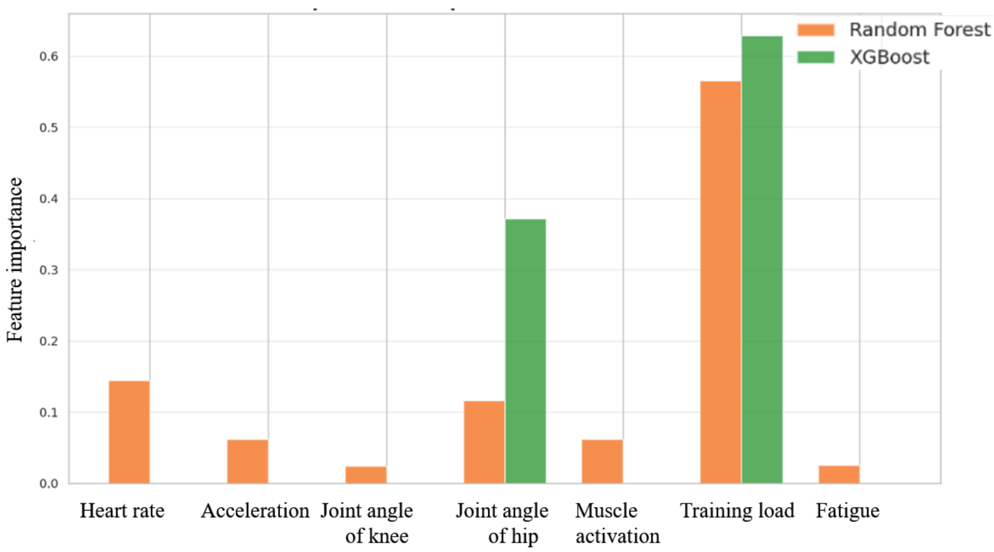

Fig. 10. (Color online) Feature importance of factors affecting injury risk.

The injury prediction models need to be optimized by focusing on the most relevant factors and refining the system architecture. In this study, training load and fatigue were significant predictors that need to be considered in the design of injury prevention programs. Training load needs to be tracked using heart rate monitors, GPS trackers, and accelerometers. Fatigue can be measured using heart rate variability and individual reporting. Reducing the number of significant predictors enhances the model's predictability, simplifies sensor requirements, and lowers costs, making injury prevention technologies affordable.

In adopting ML models, data accuracy must be ensured with the adjustment of the number of factors. Sensor accuracy fluctuations, environmental disruptions, data transmission stability, and variations in the athlete's sensor usage must also be considered to calibrate the model's performance and minimize false alarms and missed risks. Appropriate processing speed, resource efficiency, and adaptability are also important to ensure the robustness of ML models.

To monitor the training load and fatigue of athletes in predicting potential injury risks, cost-effective wearable devices can be employed to track heart rate, physical activity, and fatigue levels. Advanced systems can integrate IMUs, force plates, and environmental sensors to generate comprehensive data for elite athletic programs. A modular approach facilitates the widespread adoption of the monitoring system across diverse sports. Such a simple model needs to be tested with many athletes in diverse environmental conditions in various sports for a simple but seamless IoT-based athletic management.

5. Conclusion

We developed and validated an IoT and ML system for athlete injury prediction and performance monitoring. By leveraging real-time data from wearable sensors, the results demonstrated that training load and fatigue are critical indicators of injury risk, with advanced ensemble ML models achieving high predictive accuracy. These findings underscore the potential of the system to transform sports training by enabling personalized interventions, optimizing performance, and reducing the incidence of injuries. The ensemble models successfully identified complex patterns across diverse data sources, including heart rate variability, training intensity, and biomechanical parameters, to provide coaches and sports scientists with actionable insights. This data-driven approach offers a proactive strategy for athlete health management, supporting long-term athletic careers and consistent peak performance.

This study further demonstrates that the integration of IoT-connected real-time biomechanics with advanced ML analytics is highly effective for predicting and preventing athletic injuries. Sensor technologies, particularly IMUs and accelerometers, are essential for capturing subtle movement deviations and high-strain states that precede injury. Quantitatively, the system achieved the maximum AUC of 1.0 using ensemble models, confirming that multisensor data fusion outperforms single-factor monitoring. The data revealed that while median heart rates remained around 140 BPM, states of overexertion were identifiable through outliers exceeding 180 BPM, emphasizing the importance of continuous physiological monitoring. Ultimately, leveraging sensor technology to track variables such as joint angles, body speed, and heart rate variability enables coaches to make informed, real-time tactical adjustments. Future implementation must prioritize data privacy and cost management to ensure that these sensor-driven benefits are accessible across all levels of sports.

While the current results are promising and establish a strong foundation for future advancements, several challenges must be addressed. Data privacy and security remain paramount, requiring robust ethical frameworks and technological safeguards to protect sensitive personal information. The financial burden associated with advanced sensor technologies, data storage, and ML infrastructure presents barriers for smaller teams and amateur athletes, highlighting the need for cost-effective and scalable solutions. Additionally, resistance to adopting new technologies among those accustomed to traditional training methodologies must be mitigated through education and awareness. Refining prediction models with larger and more diverse datasets, alongside the development of accessible and affordable IoT-based monitoring systems, will be essential. Collaboration among stakeholders is critical to ensure the seamless integration of these technologies into sports practice. Addressing these challenges will enable IoT and ML systems to be effectively deployed, ensuring safer, more efficient, and more engaging training and competition environments.

References

1 Z. Chen and X. Dai: Heliyon **11** (2024) e32477. https://doi.org/10.1016/j.heliyon.2024.e32477
2 X. Liang: ACE **138** (2025) 155. https://doi.org/10.54254/2755-2721/2025.21386
3 A. Li: Int. J. Comput. Intell. Syst. **1** (2025) 23. https://doi.org/10.1007/s44196-025-00732-8
4 A, M. Rahmani, E. Yousefpoor, M. S. Yousefpoor, Z. Mehmood, A. Haider, M. Hosseinzadeh, and R. Ali Naqvi: Mathematics **22** (2021) 2970. https://doi.org/10.3390/math9222970
5 C. Cao and X. Liu: Mol. Cell. Biomech. **4** (2025) 1408. https://doi.org/10.62617/mcb1408
6 A. Hughes M. M. H. Shandhi, H. Master, J. Dunn, and E. Brittain: Circ. Res. **132** (2023) 652. https://doi.org/10.1161/CIRCRESAHA.122.322389
7 J. E. Naranjo, C. A. Mora, D. F. Bustamante Villagómez, and M. G. Mancheno Falconi, and M. V. Garcia: Sensors **25** (2025) 1526. https://doi.org/10.3390/s25051526
8 J.-Y. Wu, Y. Wang, C. T. S. Ching, H.-M. D. Wong, and L.-D. Liao: Front. Public Health **11** (2023) 1188304. https://doi.org/10.3389/fpubh.2023.1188304
9 K. Krishina Prasad, P. S. Aithal, K. G. Poormina, and K. Vinayachandra: IJHSP **5** (2021) 72. https://doi.org/10.47992/IJHSP.2581.6411.0062
10 F. Latino and F. Tafuri: Sports **12** (2024)124. https://doi.org/10.3390/sports12050124
11 H. Cho, J. Ji, Z. Chen, H. Park, and W. Lee: Procedia Comput. Sci. **52** (2015) 1083. https://doi.org/10.1016/j.procs.2015.05.119
12 A. G. Pascoal, A. Ribeiro, and J. Infante: Sports **11** (2023) 114. https://doi.org/10.3390/sports11060114
13 B. Wan and G. Shan: Springer Plus **5** (2016) 441. https://doi.org/10.1186/s40064-016-2067-y
14 The Wells Clinic: https://www.thewells-clinic.co.uk/blog/getting-to-grips-with-rsi-part-2/ (accessed June 2, 2025).
15 A. A. Jafarnezhadgero, A. Jahangirpour, H. Parsa, H. Sajedi, U. Granacher, and A. S. Souza: Sports Med. Open **9** (2023) 116. https://doi.org/10.1186/s40798-023-00663-8
16 R. Han, F. Qi, H. Wang, and M. Yi: Mol. Cell. Biomech. **3** (2024) 530. https://doi.org/10.62617/mcb530
17 M. Kolodziej, A. Groll, K. Nolte, S. Willwacher, A. Tobias, M. Schmidt, and T. Jaitner: Scand. J. Med. Sci. Sports **33** (2023) 1021. https://doi.org/10.1111/sms.14322
18 M. Calderón-Díaz, R. Silvestre Aguirre, J. P. Vásconez, R. Yáñez, M. Roby, M. Querales, and R. Salas: Sensors **24** (2024) 119. https://doi.org/10.3390/s24010119
19 J.-E. Consuegra-Fontalvo, J. Calderón-Velaides, and G.-E. Chanchí-Golondrino: Rev. Fac. Ing. **31** (2022) e204. https://www.redalyc.org/journal/4139/413973109004/movil/
20 C. Leckey, N. van Dyk, C. Doherty, A. Lawlor, and E. Delahunt: Br. J. Sports Med. **59** (2025) 491. https://doi.org/10.1136/bjsports-2024-108576

About the Authors

Jing Chen obtained her master's degree in physical education and training from Xinjiang Normal University in 2013. Since 2018, she has served as a lecturer at Shangqiu Normal University in China. Her research interests include physical education and training. (chenjing1@sqnu.edu.cn)

Wenbin Liu received his M.S. degree in computer science and technology from Fuzhou University, China. Since 2017, he has worked for Fujian New World Computer Co., Ltd., currently serving as a senior engineer. His research interests include IoT, information engineering, and sensors. He has participated in one national project, two provincial projects, and multiple horizontal industry-university research projects. He wrote textbooks, including "Software Testing Technology", "Game Testing Technology", and "Fundamentals of Digital Applications" and published more than 10 papers in domestic and international journals. (wsjy009@163.com)