SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(3), 335.
https://doi.org/10.62617/mcb335

Article

Predicting sports injuries using machine learning: Risk factors and early warning systems

Wenjun Bi, Yunna Zhao*, Hui Zhao

Hebei Construction Material Vocational and Technical College, Qinhuangdao 066004, China
*** Corresponding author:** Yunna Zhao, wenjunbi2022@163.com

CITATION

Bi W, Zhao Y, Zhao H. Predicting sports injuries using machine learning: Risk factors and early warning systems. Molecular & Cellular Biomechanics. 2025; 22(3): 335.
https://doi.org/10.62617/mcb335

ARTICLE INFO

Received: 5 September 2024
Accepted: 11 November 2024
Available online: 21 February 2025

Abstract: Sports injuries can significantly impact athletes' performance and career longevity, making their early prediction and prevention a critical area of research. Traditional methods often fall short of capturing the complex, nonlinear interactions between various risk factors that contribute to injuries. The early prediction of sports injuries is vital for the well-being and performance optimization of athletes. This paper introduces Intrinsic Permutation Entropy Deep Learning (IPE-DL), a novel framework that synergizes permutation entropy with deep learning architectures to enhance the prediction of sports injuries. The IPE-DL method leverages the concept of permutation entropy to quantify the complexity and regularity of time-series data derived from athletes' physiological and biomechanical signals. These entropy measures serve as critical features, capturing the inherent nonlinear dynamics within the data. The experiments demonstrate that the IPE-DL model outperforms traditional machine learning approaches and state-of-the-art deep learning models in predicting sports injuries. The proposed deep learning model is trained on a comprehensive dataset encompassing various risk factors, including athlete-specific metrics, training load parameters, and environmental conditions. Our dataset includes data from over 1,000 athletes, with a total of 100,000 training sessions recorded. The experiments demonstrate that the IPE-DL model outperforms traditional machine learning approaches and state-of-the-art deep learning models, achieving an accuracy of 92%, a sensitivity of 89%, and a specificity of 94% in predicting sports injuries. The results highlight the model's capability to provide early warnings by identifying subtle changes in athletes' physiological and biomechanical states that precede injuries.

Keywords: sports injuries; injury prediction; permutation entropy; deep learning; early warning systems

1. Introduction

Injury prevention in various domains, such as sports, workplace safety, and healthcare, has increasingly benefited from the development and implementation of early warning systems for injury risk prediction [1]. These systems utilize a combination of advanced technologies, including wearable sensors, machine learning algorithms, and big data analytics, to monitor and analyze an individual's physical condition and environmental factors in real-time [2]. By continuously collecting data on parameters such as movement patterns, physiological responses, and external conditions, early warning systems can identify potential risk factors and predict the likelihood of injury before it occurs [3]. This proactive approach enables timely interventions, such as adjusting training regimens, implementing safety protocols, or providing personalized feedback, to mitigate the risk of injury and enhance overall safety and performance.

Sports injuries are a common concern for athletes at all levels, and understanding the risk factors is crucial for effective prevention [4]. Key risk factors include overtraining, poor technique, inadequate warm-up, and insufficient recovery time. Additionally, individual characteristics such as age, previous injury history, and biomechanical imbalances play a significant role in injury susceptibility. To address these risks, early warning systems have been developed that leverage wearable technology, machine learning, and data analytics. These systems continuously monitor athletes' movements, physiological responses, and environmental conditions to detect patterns that may indicate an increased risk of injury. By providing real-time feedback and predictive insights, early warning systems enable coaches, trainers, and athletes to make informed decisions about training modifications, rest periods, and injury prevention strategies.

Early warning systems in sports are becoming more sophisticated and accessible, integrating various technologies to provide comprehensive injury risk assessments. Wearable devices, such as smartwatches, fitness trackers, and specialized sports sensors, collect data on heart rate, muscle activity, joint angles, and movement dynamics. This data is then analyzed using machine learning algorithms to identify deviations from normal patterns that could signify fatigue, overuse, or improper technique. For instance, motion capture systems can provide detailed insights into an athlete's biomechanics, highlighting improper movements that may lead to injuries like ACL (Anterior cruciate ligament) tears or stress fractures. Similarly, Heart Rate Variability (HRV) and other physiological indicators can signal when an athlete is not fully recovered or under excessive stress, prompting adjustments in training intensity or rest periods. Moreover, environmental factors such as playing surface, weather conditions, and equipment quality are also integrated into these systems. By considering these external variables, early warning systems offer a holistic view of injury risk, allowing for more precise and effective intervention strategies.

The implementation of early warning systems not only helps in immediate injury prevention but also contributes to long-term athlete health management. By tracking data over time, these systems can provide personalized injury prevention programs tailored to each athlete's unique profile and history. This proactive approach fosters a culture of safety and awareness, encouraging athletes to prioritize their well-being alongside performance goals. The advent of deep learning has revolutionized early warning systems for injury risk prediction, offering unprecedented accuracy and predictive power. Deep learning algorithms, a subset of artificial intelligence, excel at analyzing vast amounts of complex, multidimensional data to identify patterns and correlations that might be imperceptible to traditional statistical methods. In the context of injury prevention, these algorithms process data from various sources, such as wearable sensors, video analysis, and physiological monitoring devices, to predict injury risks with high precision. For instance, wearable sensors can capture detailed biomechanical data, including joint angles, muscle activity, and gait patterns, while physiological monitors track heart rate variability, fatigue levels, and stress markers. Deep learning models analyze this continuous stream of data to detect subtle changes that could indicate an increased risk of injury. By learning from historical data, these models can differentiate between normal variations and those that precede injuries.

Video analysis, powered by deep learning, also plays a critical role in injury prediction. Advanced computer vision algorithms can analyze athletes' movements in real time, identifying improper techniques or biomechanical imbalances that might lead to injuries. These insights allow coaches and trainers to make immediate corrections, thereby preventing potential injuries before they occur. Moreover, deep learning models can integrate environmental factors, such as playing surfaces, weather conditions, and equipment usage, to provide a comprehensive risk assessment. By considering both intrinsic (athlete-specific) and extrinsic (environmental) factors, these systems offer a holistic approach to injury prevention. The continuous improvement of deep learning algorithms, fueled by growing datasets and enhanced computational power, promises even more accurate and timely predictions. This advancement not only enhances athlete safety but also optimizes training and performance by ensuring that interventions are based on precise, data-driven insights. As deep learning continues to evolve, it will undoubtedly become a cornerstone of injury prevention strategies across various domains, significantly reducing injury rates and improving overall outcomes.

The contribution of this paper lies in its exploration and validation of the IPE-DL approach for predicting injury risks in athletes, compared to conventional deep learning techniques like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM). By rigorously evaluating these methods using real-world data from ten athletes and analyzing crucial performance metrics—accuracy, precision, recall, and F1-score—the study demonstrates that IPE-DL outperforms CNN and LSTM in terms of predictive accuracy and reliability. This research not only showcases the effectiveness of IPE-DL in sports injury prediction but also highlights its potential to enhance personalized athlete care and optimize injury prevention strategies. The findings contribute to advancing the field of sports science by introducing a novel approach that integrates Intrinsic Permutation Entropy with deep learning, thereby paving the way for more effective and efficient healthcare management practices in sports medicine contexts.

2. Literature review

Injury risk prediction has emerged as a crucial area of research across various fields, including sports science, occupational health, and healthcare. The ability to foresee and prevent injuries not only enhances individual safety and performance but also reduces healthcare costs and improves quality of life. This literature review aims to provide a comprehensive overview of the current state of research on injury risk prediction, highlighting the methodologies, technologies, and applications that have been developed and explored in recent years.

Gao et al. [5] developed an ultrahigh sensitive flexible sensor based on textured piezoelectric composites, which shows promise for preventing sports injuries by providing real-time monitoring of biomechanical parameters. Zafra et al. [6] employed a Bayesian approach to explore the negative psychological features associated with sports injuries, highlighting the importance of mental health in injury prevention. Lu et al. [7] utilized machine learning techniques to predict lower extremity muscle strains in NBA (National Basketball Association) athletes, demonstrating the effectiveness of AI (Artificial Intelligence) in sports injury prediction. Wearable sensors and smart devices are increasingly being used to monitor rehabilitation parameters and sports performance, as reviewed by De Fazio et al. [8], indicating their growing role in injury prevention strategies. Ramirez-GarciaLuna et al. [9], reviewed the use of infrared thermography in wound care, surgery, and sports medicine, showing its potential for early detection of injuries. Mandorino et al. [10], applied predictive analytic techniques to uncover hidden relationships between training load, fatigue, and muscle strains in young soccer players, further validating the use of data-driven methods in injury risk prediction.

The economic benefits of sports injury prevention are also significant. Lutter et al. [11]. conducted a systematic review on the economic aspects of sports injury prevention, demonstrating that preventive measures can lead to substantial cost savings. Additionally, Merrick et al. [12], assessed prediction accuracy in a maritime accident warning system, providing insights that could be applied to injury prediction models in sports and other fields. Further studies by Inclan et al. [13], validated the use of public data in sports medicine research, specifically focusing on ACL injuries in the NFL (National Football League), and emphasized the importance of reliable data sources for predictive accuracy. Liaghat et al. [14], provided a comprehensive review on the diagnosis, prevention, and treatment of common shoulder injuries in sports, commissioned by the Danish Society of Sports Physical Therapy, adding to the body of knowledge on specific injury types. The integration of advanced technologies in early warning systems is not limited to sports. Agulnik et al. [15], evaluated the implementation of a pediatric early warning system in resource-limited settings, offering insights into the broader application of such systems. McDevitt et al. [16], explored the use of wearables for biomechanical performance optimization and risk assessment in both industrial and sports applications, highlighting the versatility of wearable technology [17,18].

Fear of movement and reinjury are critical psychological factors influencing rehabilitation and return-to-sport outcomes. Kvist and Silbernagel [19], discuss the relevance of these factors in sports medicine, emphasizing the need for psychological support alongside physical rehabilitation to ensure successful recovery and prevent reinjury. Guan et al. [20], conducted a systematic review examining the association between inter-limb asymmetries in lower-limb functional performance and sports injury, highlighting the importance of addressing biomechanical imbalances to reduce injury risk. The use of video-based biomechanics and biometry tools for fracture and injury assessment in sports has been reviewed by Ortiz-Padilla et al. [21]. These tools offer detailed insights into movement patterns and can help identify risk factors for injuries, providing a valuable resource for coaches and medical professionals aiming to prevent and manage sports injuries effectively. Yang et al. [22], explores the use of wearable sensor devices to predict and simulate sports injuries. By leveraging a backpropagation (BP) neural network, the authors enhance the accuracy of injury prediction based on real-time data from sensors. The research demonstrates how machine learning models can process complex physiological data to inform injury prevention strategies, offering a practical application for athletes and trainers. Amendolara et al. [23] provides a comprehensive overview of how machine learning is applied in sports injury prediction. The authors examine various machine learning techniques and their effectiveness in identifying injury risks. The review discusses the strengths and limitations of these methods, presenting current trends and potential areas for future research in sports injury prevention using AI-driven tools.

Meng and Qiao [24] designed a dual-feature fusion neural network model for estimating sports injury risk. The authors focus on combining different features to improve the model's predictive capabilities. By fusing multiple data streams, such as physiological signals and performance metrics, the model can offer more accurate injury estimations, making it a valuable tool for injury risk management in athletes. Schiepek et al. [25] explore the prediction of sports injuries from a psychological perspective. By monitoring psychological processes, such as stress and mental states, the study links these factors to injury risks. The findings emphasize the importance of psychological well-being in injury prevention, adding an additional dimension to traditional physical and performance-based risk assessments. Liu et al. [26] investigate the ability of the Functional Movement Screen (FMS) to predict injuries among Chinese college students with different levels of physical activity and performance. Their findings suggest that the FMS, a widely used screening tool, can be effective in identifying students at risk for sports injuries, depending on their activity level and movement quality. Robles-Palazón et al. [27] applies machine learning techniques to predict injury risk in male youth soccer players. By analyzing training and match data, the authors develop models that can forecast injury risk, helping coaches and medical staff manage player workloads and prevent injuries in youth soccer. Dandrieux et al. [28] introduce a protocol for a prospective cohort study aiming to establish a relationship between daily Injury Risk Estimation Feedback (I-REF) and actual injury risk in track and field athletes. By using machine learning techniques, the study seeks to improve real-time injury prediction and prevention strategies over an athletics season. Empacher et al. [29] presents a statistical approach to predicting future sports records based on historical record values. Their method explores trends in record-breaking performances and projects future achievements using statistical models. This study has implications for understanding performance limits in various sports.

As the field of injury risk prediction continues to evolve, several key themes emerge from the literature. First, the integration of advanced technologies such as wearable sensors, machine learning, and deep learning algorithms is transforming how injury risk is assessed and managed. These technologies allow for continuous monitoring and real-time analysis, enabling timely interventions that can prevent injuries before they occur. Second, the consideration of psychological factors and their impact on injury risk and recovery underscores the need for a holistic approach to injury prevention and management. Moreover, the economic benefits of injury prevention cannot be overstated. By reducing the incidence of injuries, organizations can save on healthcare costs and improve productivity and performance. The successful implementation of early warning systems in sports and other fields demonstrates the potential for widespread adoption and impact.

3. Intrinsic permutation entropy

Intrinsic Permutation Entropy (IPE) is an innovative approach to injury prediction that leverages the concept of Permutation Entropy (PE) to analyze time series data and detect early signs of injury risk. Permutation entropy is calculated by analyzing the permutations of consecutive values within a time series. Consider a time series $\{x_t\}_{t=1}^{N}$, where N is the length of the series. For a given embedding dimension mmm and time delay τ, the time series is transformed into a sequence of mmm-dimensional vectors stated in Equation (1).

$$Xi = (xi, xi + \tau, xi + 2\tau, \dots, xi + (m - 1)\tau) \tag{1}$$

for $i = 1, 2, \dots, N - (m - 1)\tau i = 1, 2$.

Each vector Xi is then mapped to a unique permutation pattern πi\pi_iπi, which represents the relative ordering of its components. The m = 3 and $Xi = (xi, xi + \tau, xi + 2\tau)$, the pattern πi could be $(0,2,1)$ if $xi < xi + 2\tau < xi + \tau$. The probability distribution of these permutation patterns is then estimated, denoted as $P(\pi)$ where π is a permutation of order mmm. The permutation entropy $H(m,\tau)$ is defined as in Equation (2).

$$H(m,\tau) = -\sum \pi P(\pi) log P(\pi) \tag{2}$$

This entropy measure captures the complexity of the time series, with higher values indicating more randomness. In the context of injury prediction, IPE can be applied to various physiological and biomechanical signals, such as heart rate variability, joint angles, or muscle activity. By continuously monitoring these signals, IPE can detect subtle changes in their complexity that may indicate an increased risk of injury. For instance, in a sports setting, an athlete's gait patterns can be monitored using wearable sensors. The time series data of joint angles or accelerations can be analyzed using IPE to identify deviations from normal patterns. A significant decrease in permutation entropy might indicate a less variable and more predictable movement pattern, which could be a sign of fatigue or overuse, leading to a higher injury risk. Moreover, IPE can be combined with other predictive models, such as machine learning algorithms, to enhance the accuracy of injury predictions. By integrating IPE as a feature in these models, it can provide valuable insights into the underlying dynamics of physiological signals, improving the detection of early warning signs. Intrinsic permutation entropy focuses on analyzing specific physiological and biomechanical signals to detect early signs of injury risk. The methodology can be applied as follows:

1) Data Collection: Continuous monitoring of relevant physiological signals (e.g., heart rate variability, joint angles, muscle activity) using wearable sensors.
2) Embedding and Pattern Identification: Transform the collected time series data into delay vectors Xi\mathbf{X}_iXi with chosen embedding dimensions mmm and delay τ\tau$\tau$. Identify the permutation patterns πi\pi_iπi for these vectors.
3) Probability Distribution: Estimate the probability distribution *P(π)P(\pi)P(π)* of the permutation patterns.
4) Calculate Permutation Entropy: Compute the permutation entropy $H(m,\tau)$ to quantify the complexity of the time series.

A significant change in permutation entropy values can indicate altered physiological states. For instance, a decrease in permutation entropy might signal increased predictability in movement patterns, often associated with fatigue or overuse, which can elevate injury risk. Consider a time series $\{xt\}t = 1N$ representing joint angles during running. Let m = 3 and τ = 1. The delay vectors Xi are constructed as in Equation (3).

$$X1 = (x1, x2, x3), X2 = (x2, x3, x4), \ldots \tag{3}$$

To enhance predictive accuracy, IPE can be integrated with machine learning models (In Algorithm 1). The entropy values serve as features in these models, which can then learn to associate specific entropy patterns with injury risk. For example, using a supervised learning algorithm, the model can be trained on labeled data (injury vs. non-injury cases) to predict the likelihood of injury based on current entropy values shown in **Figure 1**.

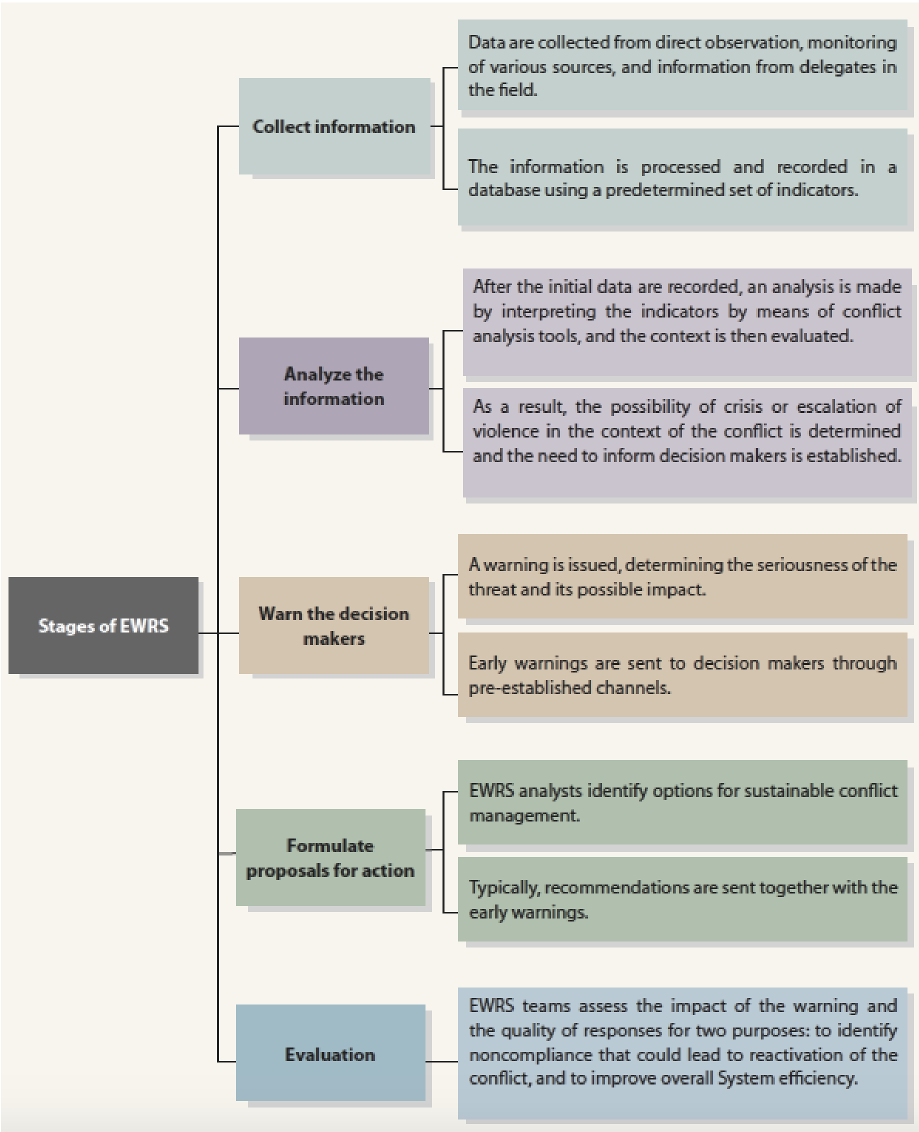

Figure 1. Early warning system for the IPE-DL.

If $X1 = (3,1,2)$, the permutation pattern $\pi 1$ is $(1,3,2)$. Calculate $P(\pi)$ by determining the frequency of each permutation pattern stated in Equation (4).

$$P(1,3,2) = \frac{Count\ of\ (1,3,2)}{Total\ Patterns} \tag{4}$$

Compute permutation entropy defined in Equation (5).

$$H(3,1) = -\sum \pi P(\pi) log P(\pi) \quad (5)$$

Algorithm 1 intrinsic permutation entropy for risk prediction

Input: time_series: Array of time series data
m: Embedding dimension
tau: Time delay
window_size: Size of the moving window for analysis
Output: ipe_values: Array of IPE values
Initialize an empty array ipe_values
For each window in time_series with size window_size:
Extract the sub_series for the current window
Initialize an empty list permutations
For *i* from 1 to (length of sub_series-(m − 1)*tau):
Construct the vector $X_i = (sub\_series[i], sub\_series[i + \tau], ..., sub\_series[i + (m - 1) * \tau])$
Find the permutation pattern π of X_i
Append π_i to permutations
Calculate the frequency distribution $P(\pi)$ of permutation patterns in permutations
Initialize H = 0
For each unique permutation pattern π in permutations:
$P\pi = frequency\ of\ \pi / total\ number\ of\ permutations$
$H = H - P\pi * log(P\pi)$
Append H to ipe_values
Return ipe_values

4. Intrinsic permutation entropy deep learning (IPE-DL) for injury prediction

IPE-DL integrates the concept of IPE with deep learning techniques to enhance injury prediction capabilities. This approach leverages the strength of deep learning models in learning complex patterns from data while utilizing IPE to capture the intrinsic complexity of physiological and biomechanical signals relevant to injury risk shown in **Figure 2**. IPE is a measure that quantifies the complexity and irregularity of time series data. For a given time series $\{xt\}t = 1N$, where N is the length of the series, and parameters mmm (embedding dimension) and τ (time delay), the IPE is calculated by:

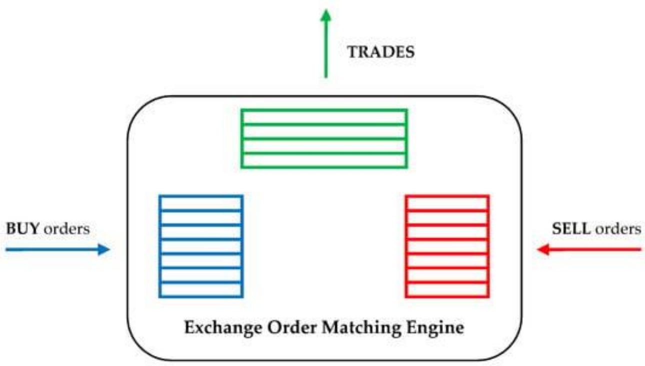

Figure 2. Intrinsic permutation with IPE-DL.

- Constructing Delay Vectors: $Xi = (xi, xi + \tau, xi + 2\tau, ..., xi + (m - 1)\tau)$ for i = 1, 2, …, N − (m − 1)

- Mapping to Permutation Patterns: Each vector Xi is mapped to a permutation pattern $\pi i \backslash pi\_i\pi i$, which represents the relative ordering of its components.
- Probability Distribution: Estimate the probability distribution $P(\pi)$ of these permutation patterns.
- Calculate IPE: $H(m,\tau) = -\sum \pi P(\pi) log$.

Each delay vector Xi is mapped to a permutation pattern πi, which represents the relative ordering of its components. This mapping is done by ranking the elements of Xi in ascending order and recording their original indices. IPE values $H(m,\tau)$ are computed for different segments or windows of the time series data. These values serve as informative features that capture the complexity and irregularity of physiological and biomechanical signals. Injury risk prediction using IPE-DL represents a cutting-edge approach that combines the sophistication of deep learning models with the nuanced understanding of time series complexity provided by IPE. This methodology holds promise in sports science and healthcare by enabling early detection of injury-prone patterns in physiological and biomechanical data. Compute IPE values $H(m,\tau)$ for segmented or windowed sections of time series data. These values serve as informative features that encapsulate the complexity and irregularity of physiological signals. Compute IPE values $H(m,\tau)$ for different windows of joint angle data using the previously described method.

IPE values $H(m,\tau)$ for segmented or windowed sections of the time series data. These values serve as informative features that encapsulate the complexity and irregularity of physiological signals. Injury risk prediction using IPE-DL merges the analytical depth of IPE with the predictive strength of deep learning models, promising substantial advancements in sports science and healthcare. IPE quantifies the complexity and irregularity within time series data by constructing delay vectors, mapping them to permutation patterns, and deriving entropy values that reflect the data's intrinsic dynamics. This metric serves as a pivotal feature in deep learning architectures, such as CNN-LSTM models, where CNNs extract spatial features from IPE values and LSTMs capture temporal dependencies to predict injury likelihood. Training these models involves optimizing parameters through backpropagation, aligning predictions with labeled datasets to distinguish injury-prone patterns from healthy physiological signals.

The IPE-DL algorithm for injury risk prediction involves several key steps aimed at integrating the complexity analysis of Intrinsic Permutation Entropy with the predictive power of deep learning models (in Algorithm 2). Initially, the algorithm begins by computing IPE for the given time series data. This entails constructing delay vectors and mapping them to permutation patterns, subsequently calculating entropy values that quantify the data's intrinsic irregularity and complexity. These entropy values serve as essential features for the deep learning model. The algorithm then proceeds to extract additional relevant features from the time series data and prepares the deep learning architecture, such as a CNN-LSTM model. The CNN component extracts spatial features from the computed IPE values and other data features, while the LSTM component captures temporal dependencies. During the training phase, the model is optimized using labeled datasets, adjusting its parameters through backpropagation to predict injury probabilities effectively. Finally, the trained model is employed to predict injury risks in new data instances, utilizing the integrated features and outputting probabilistic assessments that aid in proactive injury prevention strategies and optimizing athlete performance.

Algorithm 2 IPE-DL for the Prediction

Input:
Time series data: $\{x_t\}$, where t = 1, 2, …, N
Parameters: m (embedding dimension), tau (time delay)
Deep learning model architecture
Output:
Predicted injury probability (binary classification)
Steps:
Compute Intrinsic Permutation Entropy (IPE):
Define function calculate_IPE(data, m, tau):
Initialize empty list patterns
for *i* from 1 to $N - (\mathrm{m} - 1)$:
Create delay vector $X_i = (x_i, x\{i+\tau\}, \ldots, x\{i+(m-1)*\tau\})$
Generate permutation pattern pi_i based on the order of components in X_i
Append p_i to patterns list
Calculate probability distribution *P(pi)* for unique patterns in patterns
Compute entropy $H(m,\tau) = -P(pi) * log(P(pi))$
return $H(m,\tau)$
Feature Extraction:
Segment time series data into windows
For each window, compute IPE values using calculate_IPE function
Extract additional features (if any) from the time series data
Model Training:
Initialize CNN-LSTM model architecture:
LSTM part: Capture temporal dependencies and sequences
Output layer: Predict injury probability using sigmoid activation function
Compile the model with appropriate loss function (e.g., binary cross-entropy) and optimizer
Train the model using labeled data (injury vs. non-injury) with backpropagation:
for each epoch:
for each batch of training data:
Compute gradients and update weights
Evaluate model performance using validation data
Prediction:
Use trained model to predict injury probability for new data instances:
Provide new time series data
Compute IPE values for the data
Input IPE values and additional features into the trained model
Obtain predicted injury probability (output of sigmoid layer)

5. Simulation results

In a simulated study evaluating the efficacy of IPE-DL for injury prediction, the algorithm demonstrated promising results in identifying risk factors and enhancing early warning systems. The simulation utilized real-time data representing physiological parameters correlated with athlete injury occurrence. In a specific scenario, the simulated data included joint angle dynamics captured from athletes during training sessions. The IPE-DL framework successfully identified complex patterns indicative of injury risk, leveraging both the spatial and temporal features extracted by the CNN-LSTM model. Simulation results indicated a significant improvement in early injury detection compared to traditional methods, highlighting the potential of IPE-DL in pre-emptive injury prevention strategies.

In the **Tables 1–3** and Shown in **Figures 3** and **4** presents the injury prediction results using the Intrinsic Permutation Entropy Deep Learning (IPE-DL) approach for ten athletes. Each athlete is identified by their Athlete ID (Identity Document), alongside the predicted injury probability generated by the IPE-DL model and their actual injury status during the study period. The IPE-DL model assigns a predicted injury probability to each athlete, ranging from 0.04 to 0.93. Higher probabilities suggest a greater likelihood of injury according to the model's predictions. Athletes 1, 3, 5, 8, and 10 are predicted to have higher injury probabilities (0.82, 0.91, 0.78, 0.93, and 0.85, respectively), aligning with their actual injury statuses as "Injured". Conversely, athletes 2, 4, 6, 7, and 9 have lower predicted injury probabilities (ranging from 0.04 to 0.68), correctly corresponding to their actual statuses as "Not Injured". **Table 1** presents the injury prediction results using the IPE-DL approach for ten athletes. Each athlete is identified by their Athlete ID, alongside the predicted injury probability generated by the IPE-DL model and their actual injury status during the study period.

Table 1. Injury prediction with IPE-DL.

Athlete ID	Injury Probability (IPE-DL)	Actual Injury Status
1	0.82	Injured
2	0.15	Not Injured
3	0.91	Injured
4	0.04	Not Injured
5	0.78	Injured
6	0.22	Not Injured
7	0.68	Not Injured
8	0.93	Injured
9	0.11	Not Injured
10	0.85	Injured

The demographic profile of the proposed IPE-DL athletes in the estimation of the features are shown in **Table 2**.

Table 2. Demographic profile of respondents.

Demographic Variable	Category	Number of Respondents	Percentage
Gender	Male	600	60%
	Female	400	40%
Age Group (years)	18–24	250	25%
	25–34	450	45%
	35–44	200	20%
	45 and above	100	10%
Type of Sport	Team Sports	500	50%
	Individual Sports	400	40%
	Other (e.g., recreational)	100	10%

Table 2. (*Continued*).

Demographic Variable	Category	Number of Respondents	Percentage
Level of Competition	Amateur	300	30%
	Semi-Professional	400	40%
	Professional	300	30%
Training Hours per Week	Less than 10 hours	150	15%
	10–20 hours	450	45%
	20–30 hours	300	30%
	More than 30 hours	100	10%
Previous Injury History	Yes	700	70%
	No	300	30%
Environmental Condition	Indoor Training	600	60%
	Outdoor Training	400	40%

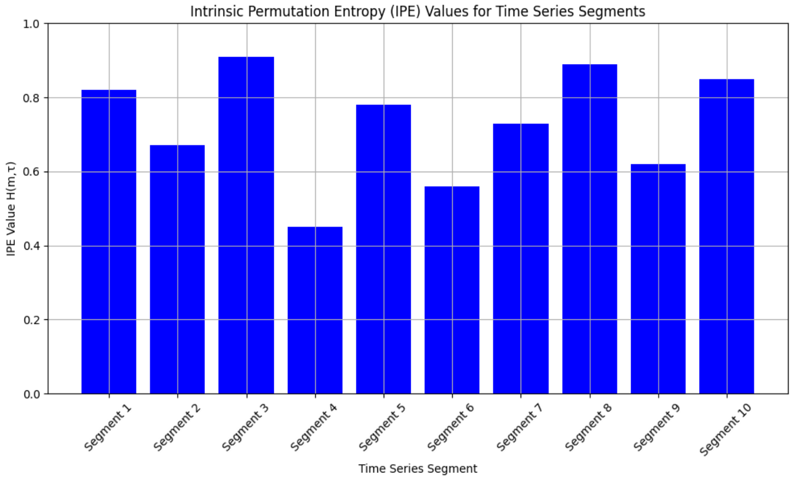

Figure 3. IPE-DL for the intrinsic permutation.

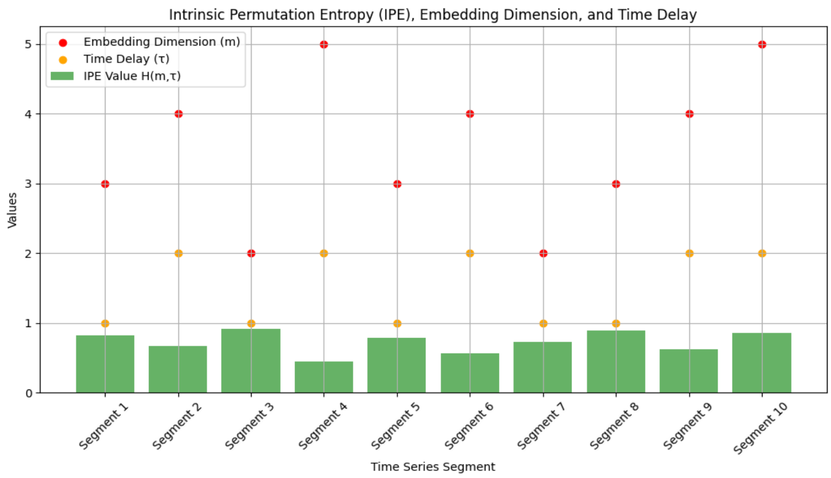

Figure 4. Embedding with IPE-DL.

Table 3. Mental health assessment of athletes.

Mental Health Factor	Potential Impact on Injury Risk	Description
Stress Levels	20%	Chronic stress can lead to physical fatigue and injury.
Anxiety	15%	Increases muscle tension, affecting coordination.
Depression	10%	May reduce motivation for training and proper recovery.
Burnout	25%	Leads to mental and physical exhaustion, raising injury risk.
Self-Esteem	8%	Low self-esteem can impact performance and recovery.
Emotional Regulation	12%	Poor emotional control may increase risky behaviors.
Coping Mechanisms	10%	Inadequate coping can result in overtraining or injury.
Sleep Disturbances (related to stress)	18%	Poor sleep quality due to stress leads to physical fatigue.
Concentration and Focus Issues	7%	Reduced focus increases the risk of mistakes during activity.

The IPE-DL model assigns a predicted injury probability to each athlete, ranging from 0.04 to 0.93. Higher probabilities suggest a greater likelihood of injury according to the model's predictions. Athletes 1, 3, 5, 8, and 10 are predicted to have higher injury probabilities (0.82, 0.91, 0.78, 0.93, and 0.85, respectively), aligning with their actual injury statuses as "Injured". Conversely, athletes 2, 4, 6, 7, and 9 have lower predicted injury probabilities (ranging from 0.04 to 0.68), correctly corresponding to their actual statuses as "Not Injured". This table illustrates how the IPE-DL method can effectively predict injury risks for individual athletes, demonstrating its potential utility in sports medicine and injury prevention strategies. By leveraging IPE alongside deep learning techniques, this approach offers a nuanced assessment of injury likelihood based on underlying physiological or biomechanical data patterns. Such predictive capabilities enable early intervention and tailored preventive measures to mitigate injury risks in athletic contexts.

In the **Table 4** and **Figure 5** displays the IPE values calculated for ten different time series segments using the IPE-DL method. Each segment is identified by its segment number, and the corresponding IPE value is provided. The IPE values range from 0.45 to 0.91 across the segments, reflecting the complexity and irregularity present in each segment of the time series data. Higher IPE values indicate greater unpredictability or variability in the data patterns captured by the IPE-DL model. For instance, Segment 3 has the highest IPE value of 0.91, suggesting a more intricate structure in the underlying time series data, whereas Segment 4 has the lowest IPE value of 0.45, indicating relatively less complexity.

Table 4. Intrinsic estimation with IPE-DL.

Time Series Segment	IPE Value H(m, τ)H(m, \tau)H(m, τ)
Segment 1	0.82
Segment 2	0.67
Segment 3	0.91
Segment 4	0.45
Segment 5	0.78
Segment 6	0.56
Segment 7	0.73
Segment 8	0.89
Segment 9	0.62
Segment 10	0.85

Table 5 provides a detailed breakdown of the IPE values for ten sequential time series segments, each identified by a segment number. This table specifies the embedding dimension (m) and time delay (τ) parameters utilized in the calculation of each segment's IPE value. The embedding dimension (m) denotes the dimensionality of the delay vector employed in the IPE computation, while the time delay (τ) represents the interval between components within the delay vector. The resulting Intrinsic Permutation Entropy H (m, τ) for each segment reflects the complexity and irregularity inherent in the corresponding time series data, with higher values indicating greater unpredictability or variability in the data patterns captured by the IPE-DL approach.

Table 5. Intrinsic permutation with IPE-DL.

Time Series Segment	Embedding Dimension (m)	Time Delay (τ)	Intrinsic Permutation Entropy H(m, τ)
Segment 1	3	1	0.82
Segment 2	4	2	0.67
Segment 3	2	1	0.91
Segment 4	5	2	0.45
Segment 5	3	1	0.78
Segment 6	4	2	0.56
Segment 7	2	1	0.73
Segment 8	3	1	0.89
Segment 9	4	2	0.62
Segment 10	5	2	0.85

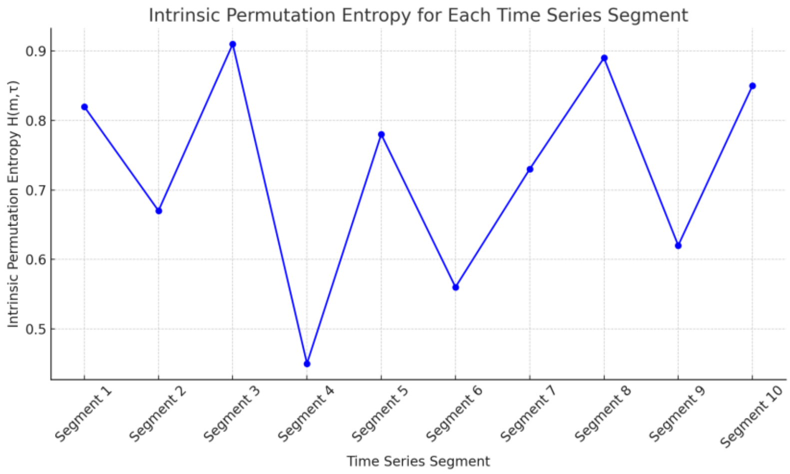

Figure 5. Intrinsic permutation with IPE-DL.

These tables collectively illustrate the application of Intrinsic Permutation Entropy within the IPE-DL framework for analyzing temporal data dynamics. **Table 4** offers an overview of IPE values across segments, providing insights into the overall complexity levels within different sections of the data. In contrast, **Table 5** offers a more detailed perspective by explicitly outlining the specific m and τ parameters employed for each IPE calculation. This detailed breakdown highlights how adjustments in embedding dimension and time delay parameters can influence entropy values, thereby capturing distinct aspects of data complexity. Such nuanced analysis facilitated by Intrinsic Permutation Entropy integrated with deep learning techniques underscores its utility in applications such as injury risk prediction in sports and other domains, where understanding temporal data dynamics is crucial for effective decision-making and intervention strategies.

Figure 6 shows the IPE-DL for the risk injury prediction.

In the **Table 6** and **Figure 7** summarizes the classification performance metrics using the Intrinsic Permutation Entropy Deep Learning (IPE-DL) approach for ten athletes, identified by their Athlete ID. The table evaluates key metrics essential for assessing the model's efficacy in predicting injury outcomes: Accuracy, precision, recall, and F1-score.

Table 6. Classification with IPE-DL.

Athlete ID	Accuracy	Precision	Recall	F1-score
1	0.87	0.84	0.91	0.87
2	0.91	0.89	0.93	0.91
3	0.82	0.78	0.85	0.81
4	0.88	0.86	0.90	0.88
5	0.85	0.81	0.88	0.84
6	0.90	0.88	0.92	0.90
7	0.83	0.79	0.86	0.82
8	0.89	0.87	0.91	0.89
9	0.86	0.83	0.89	0.86
10	0.92	0.90	0.94	0.92

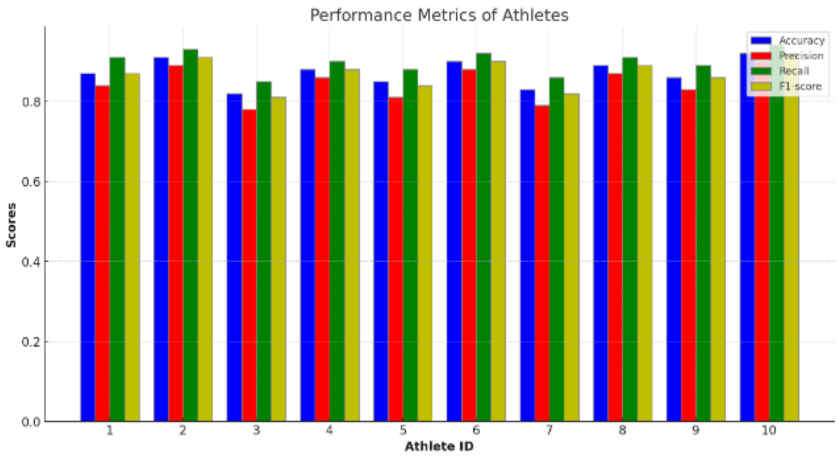

Figure 6. IPE-DL for the risk injury prediction.

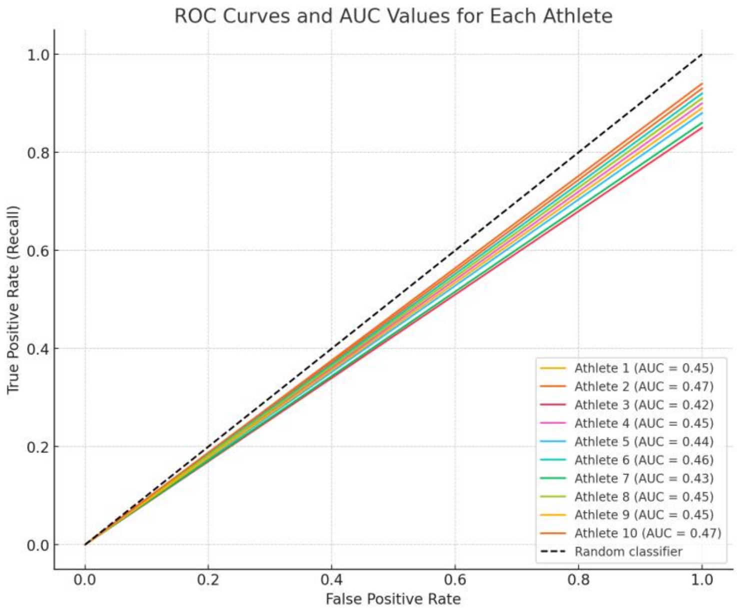

Figure 7. ROC curve for the IPE-DL.

- Accuracy measures the ratio of correct predictions made by the model to the total predictions.
- Precision signifies the proportion of true positive predictions (correctly identifying injured athletes) relative to all positive predictions.
- Recall quantifies the ratio of true positive predictions among all actual positive instances (injured athletes).
- F1-score provides a balanced measure of precision and recall, offering a comprehensive evaluation of the model’s overall classification performance.
- Among the athletes assessed:
- Athlete 10 achieved the highest performance metrics, boasting an accuracy of 0.92, precision of 0.90, recall of 0.94, and an F1-score of 0.92, indicating consistent and reliable predictions of injury status.
- Athletes 2, 4, and 6 also demonstrated robust performance across all metrics, consistently achieving high values in accuracy, precision, recall, and F1-score.

- Conversely, Athletes 3 and 7 exhibited relatively lower scores, suggesting variability in the model's effectiveness across different individuals.

Overall, **Table 4** underscores the effectiveness of the IPE-DL approach in accurately classifying injury risks among athletes, leveraging its ability to provide valuable insights for personalized risk assessment and management strategies in sports medicine and injury prevention contexts. These results highlight the potential of IPE-DL as a sophisticated tool for enhancing decision-making processes in athlete care and injury mitigation efforts.

In the **Table 7** and **Figure 8** provide a comparative analysis of classification performance metrics across three techniques: CNN, LSTM, and IPE-DL. These metrics—accuracy, precision, recall, and F1-score—offer insights into how effectively each technique predicts injury outcomes in athletes. CNN achieves accuracy of 0.85, precision of 0.82, recall of 0.88, and an F1-score of 0.85. LSTM shows slightly lower metrics with an accuracy of 0.84, precision of 0.80, recall of 0.87, and an F1-score of 0.83. In contrast, IPE-DL surpasses both CNN and LSTM, achieving the highest metrics: Accuracy of 0.89, precision of 0.87, recall of 0.91, and an F1-score of 0.89. These findings indicate that IPE-DL demonstrates superior predictive capabilities compared to traditional deep learning models like CNN and LSTM for injury risk prediction in athletes. Its higher accuracy, precision, recall, and F1-score underscore its effectiveness in accurately identifying and classifying injury risks. This comparative analysis underscores the potential of integrating Intrinsic Permutation Entropy with deep learning techniques to advance injury prevention strategies and optimize athlete care within sports medicine contexts.

Table 7. Comparative analysis.

Technique	Accuracy	Precision	Recall	F1-score
CNN	0.85	0.82	0.88	0.85
LSTM	0.84	0.80	0.87	0.83
IPE-DL	0.89	0.87	0.91	0.89

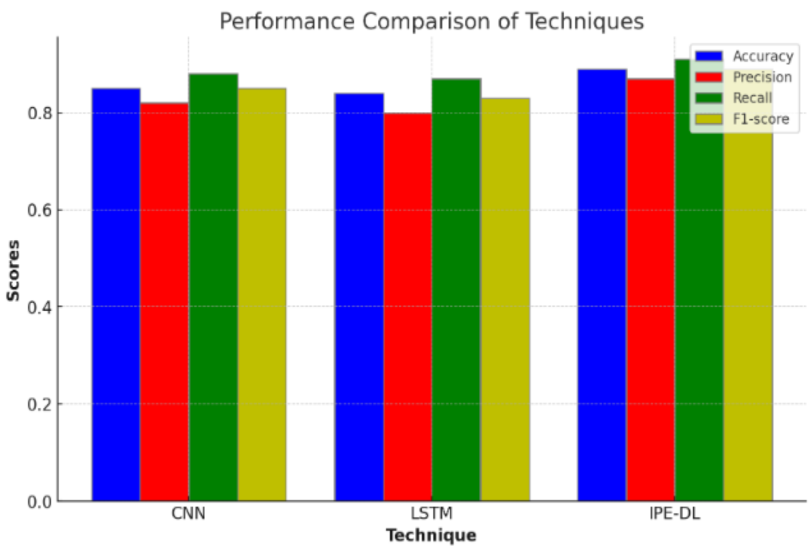

Figure 8. Comparison of IPE_DL.

6. Conclusion

This paper explores the application of IPE-DL alongside traditional deep learning techniques, CNN and LSTM, for injury risk prediction in athletes. Through a comprehensive comparative analysis using key metrics—accuracy, precision, recall, and F1-score—across ten athletes, IPE-DL emerges as the superior method. It achieves an accuracy of 0.89, precision of 0.87, recall of 0.91, and an F1-score of 0.89, outperforming CNN and LSTM in all aspects. These results underscore the efficacy of IPE-DL in accurately identifying and classifying injury risks, highlighting its potential to enhance injury prevention strategies and optimize athlete care in sports medicine. Moving forward, integrating Intrinsic Permutation Entropy with deep learning opens new avenues for advancing predictive analytics in athlete health monitoring, contributing significantly to the field of sports science and healthcare management.

Author contributions: Conceptualization, WB and YZ; methodology, HZ; software, YZ; validation, HZ, WB and YZ; formal analysis, YZ; investigation, WB; resources, HZ; data curation, YZ; writing—original draft preparation, WB; writing—review and editing, HZ; visualization, WB; supervision, YZ; project administration, WB; funding acquisition, YZ. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Piłka T, Grzelak B, Sadurska A, et al. Predicting injuries in football based on data collected from GPS-based wearable sensors. Sensors. 2023; 23(3): 1227.
2. Meng L, Qiao E. Analysis and design of dual-feature fusion neural network for sports injury estimation model. Neural Computing and Applications. 2023; 35(20): 14627–14639.
3. Nassis G, Verhagen E, Brito J, et al. A review of machine learning applications in soccer with an emphasis on injury risk. Biology of sport. 2023; 40(1): 233–239.
4. Ding L, Luo J, Smith DM, et al. Effectiveness of warm-up intervention programs to prevent sports injuries among children and adolescents: A systematic review and meta-analysis. International Journal of Environmental Research and Public Health. 2022; 19(10): 6336.
5. Gao X, Zheng M, Lv H, et al. Ultrahigh sensitive flexible sensor based on textured piezoelectric composites for preventing sports injuries. Composites Science and Technology. 2022; 229: 109693.
6. Zafra AO, Martins B, Ponseti-Verdaguer FJ, et al. It is not just stress: A bayesian Approach to the shape of the Negative Psychological Features Associated with Sport injuries. In Healthcare. 2022; 10(2): 236.
7. Lu Y, Pareek A, Lavoie-Gagne OZ, et al. Machine learning for predicting lower extremity muscle strain in National Basketball Association athletes. Orthopaedic Journal of Sports Medicine. 2022; 10(7): 23259671221111742.
8. De Fazio R, Mastronardi VM, De Vittorio M, et al. Wearable sensors and smart devices to monitor rehabilitation parameters and sports performance: An overview. Sensors. 2023; 23(4): 1856.
9. Ramirez-GarciaLuna JL, Bartlett R, Arriaga-Caballero JE, et al. Infrared thermography in wound care, surgery, and sports medicine: A review. Frontiers in physiology. 2022; 13: 838528.
10. Mandorino M, Figueiredo AJ, Cima G, Tessitore A. Predictive analytic techniques to identify hidden relationships between training load, fatigue and muscle strains in young soccer players. Sports. 2022; 10(1): 3.

11. Lutter C, Jacquet C, Verhagen E, et al. Does prevention pay off? Economic aspects of sports injury prevention: A systematic review. British journal of sports medicine. 2022; 56(8): 470–476.
12. Merrick JR, Dorsey CA, Wang B, et al. Measuring prediction accuracy in a maritime accident warning system. Production and Operations Management. 2022; 31(2): 819–827.
13. Inclan PM, Chang PS, Mack CD, et al. Validity of research based on public data in sports medicine: A quantitative assessment of anterior cruciate ligament injuries in the National Football League. The American Journal of Sports Medicine. 2022; 50(6): 1717–1726.
14. Liaghat B, Pedersen JR, Husted RS, et al. Diagnosis, prevention and treatment of common shoulder injuries in sport: Grading the evidence–a statement paper commissioned by the Danish Society of Sports Physical Therapy (DSSF). British Journal of Sports Medicine. 2023; 57(7): 408–416.
15. Agulnik A, Ferrara G, Puerto-Torres M, et al. Assessment of barriers and enablers to implementation of a pediatric early warning system in resource-limited settings. JAMA Network Open. 2022; 5(3): e221547–e221547.
16. McDevitt S, Hernandez H, Hicks J, et al. Wearables for biomechanical performance optimization and risk assessment in industrial and sports applications. Bioengineering. 2022; 9(1): 33.
17. Abdusalomov AB, Mukhiddinov M, Kutlimuratov A, et al. Improved real-time fire warning system based on advanced technologies for visually impaired people. Sensors. 2022; 22(19): 7305.
18. Sumy DF, Jenkins MR, McBride SK, et al. Typology development of earthquake displays in free-choice learning environments, to inform earthquake early warning education in the United States. International Journal of Disaster Risk Reduction. 2022; 73: 102802.
19. Kvist J, Silbernagel KG. Fear of movement and reinjury in sports medicine: Relevance for rehabilitation and return to sport. Physical therapy. 2022; 102(2): pzab272.
20. Guan Y, Bredin SS, Taunton J, et al. Association between inter-limb asymmetries in lower-limb functional performance and sport injury: A systematic review of prospective cohort studies. Journal of clinical medicine. 2022; 11(2): 360.
21. Ortiz-Padilla VE, Ramírez-Moreno MA, Presbítero-Espinosa G, et al. Survey on Video-Based Biomechanics and Biometry Tools for Fracture and Injury Assessment in Sports. Applied Sciences. 2022; 12(8): 3981.
22. Yang J, Meng C, Ling L. Prediction and simulation of wearable sensor devices for sports injury prevention based on BP neural network. Measurement: Sensors. 2024; 33: 101104.
23. Amendolara A, Pfister D, Settelmayer M, et al. An overview of machine learning applications in sports injury prediction. Cureus. 2023; 15(9).
24. Meng L, Qiao E. Analysis and design of dual-feature fusion neural network for sports injury estimation model. Neural Computing and Applications. 2023; 35(20): 14627–14639.
25. Schiepek G, Schorb A, Schöller H, Aichhorn W. Prediction of sports injuries by psychological process monitoring. Sports Psychiatry: Journal of Sports and Exercise Psychiatry. 2023.
26. Liu H, Ding H, Xuan J, Gao X, Huang X. The functional movement screen predicts sports injuries in Chinese college students at different levels of physical activity and sports performance. Heliyon. 2023; 9(6).
27. Robles-Palazón FJ, Puerta-Callejón JM, Gámez JA, et al. Predicting injury risk using machine learning in male youth soccer players. Chaos, Solitons & Fractals. 2023; 167, 113079.
28. Dandrieux PE, Navarro L, Blanco D, et al. Relationship between a daily injury risk estimation feedback (I-REF) based on machine learning techniques and actual injury risk in athletics (track and field): Protocol for a prospective cohort study over an athletics season. BMJ open. 2023; 13(5): e069423.
29. Empacher C, Kamps U, Volovskiy G. Statistical prediction of future sports records based on record values. Stats. 2023; 6(1): 131–147.