Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

ARTIFICIAL INTELLIGENCE–DRIVEN SPORTS SCIENCE INNOVATION IN SWITZERLAND: A MACHINE LEARNING–BASED SMART FRAMEWORK FOR PERSONALIZED TRAINING, ATHLETIC PERFORMANCE OPTIMIZATION, AND INJURY RISK REDUCTION

Francesco Ernesto Alessi Longa[1], Dr. Maira Anis[*2], Syed Ahmed Raza[3], Shoaib Abbas[4]

[1]*Department of Sport Sciences, Unitelematica Leonardo da Vinci, Switzerland.*
[*2]*NUST Business School, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan.*
[3]*Department of Computer Science, SZABIST University Hyderabad Campus, Pakistan.*
[4]*School of Computer and Information Science, University of the Cumberlands, Kentucky, United State of America*

[1]ceskone@libero.it, [*2]maira.anis@nbs.nust.edu.pk, [3]ahmed.raza@hyd.szabist.edu.pk, [4]smohammad16535@ucumberlands.edu

[1]**ORCID:** 0009-0002-6068-6203

DOI: https://doi.org/10.5281/zenodo.21817893

Keywords
Artificial intelligence; Machine learning; Sports science; Personalized training; Athletic performance optimization; Injury risk prediction; Wearable sensors; Switzerland.

Article History
Received: 18 January 2026
Accepted: 03 March 2026
Published: 17 March 2026

Corresponding Author: *
Dr. Maira Anis

Abstract
Artificial intelligence is reshaping sports science by enabling continuous athlete monitoring, personalized training prescription, performance forecasting, and early injury-risk detection. This study developed an artificial intelligence–driven smart framework for Swiss sports organizations by integrating machine learning, wearable-sensor data, physiological indicators, training-load records, and injury histories. A multisource dataset was constructed from 320 competitive athletes representing football, athletics, cycling, skiing, and indoor team sports across Swiss university and regional training centers. Data were collected over 24 weeks from GPS devices, heart-rate monitors, inertial measurement units, athlete wellness questionnaires, medical screening records, and performance tests. The final dataset contained 18,740 athlete-session observations and 42 variables, including age, sex, sport type, training duration, total distance, sprint count, acceleration load, heart-rate variability, resting heart rate, sleep duration, perceived fatigue, recovery score, previous injury, body mass index, jump performance, and session rating of perceived exertion. After data cleaning, normalization, imbalance correction, and recursive feature selection, random forest, support vector machine, XGBoost, and artificial neural network models were trained using stratified ten-fold cross-validation. XGBoost achieved the best injury-risk classification performance, with 92.6% accuracy, 0.91 precision, 0.89 recall, a 0.90 F1-score, and an area under the receiver operating characteristic curve of 0.95. The neural network produced the strongest performance-prediction results, achieving an R^2 of 0.86, a root mean square error of 0.12, and a mean absolute error of 0.08. Compared with conventional training plans, the proposed framework improved predicted performance by 14.8%, reduced excessive workload

exposure by 21.3%, and lowered estimated injury risk by 27.6%. Feature-importance analysis identified acute-to-chronic workload ratio, previous injury, heart-rate variability, sleep duration, and acceleration load as the most influential predictors. The findings indicate that an AI-supported, athlete-centered framework can enhance training individualization, improve performance outcomes, and support preventive sports medicine in Switzerland. The framework also incorporates model transparency, secure data governance, and privacy-preserving decision support, providing a scalable foundation for coaches, clinicians, and sports institutions within both elite and developmental sporting environments.

1- Introduction:

The rapid development of artificial intelligence has created new opportunities for transforming sports science, athlete monitoring, training design, performance evaluation, and injury prevention. Traditional sports-training programs have generally relied on standardized schedules, periodic physical assessments, coaches' observations, and retrospective medical evaluations. Although these approaches remain valuable, they may not fully capture the complex and continuously changing physiological, biomechanical, psychological, and environmental conditions that influence athletic performance. Athletes respond differently to similar training loads because of variations in age, fitness level, previous injury, recovery capacity, sleep quality, nutrition, stress, and sport-specific demands. Consequently, there is an increasing need for intelligent systems capable of analyzing multidimensional athlete data and translating it into individualized, timely, and evidence-based training decisions. Artificial intelligence, particularly machine learning, provides a powerful foundation for addressing this challenge [1]. Machine-learning algorithms can identify hidden patterns within large sports datasets, estimate future performance, classify injury risk, detect abnormal workload responses, and support personalized training recommendations. When combined with wearable sensors, global positioning systems, heart-rate monitors, inertial measurement units, wellness questionnaires, and medical records, these models can provide a more comprehensive representation of an athlete's physical condition. Such systems may help coaches and sports scientists move from generalized training prescriptions toward adaptive and athlete-centered interventions. Personalized training is especially important in high-performance sport because the effectiveness of a training program depends on the balance between workload, adaptation, fatigue, and recovery. Insufficient training may limit performance development, whereas excessive or poorly managed training may result in overtraining, reduced performance, illness, or injury [2]. Conventional workload-monitoring methods often assess only a limited number of indicators and may not recognize nonlinear relationships between internal and external training loads. Machine-learning models can simultaneously process variables such as training duration, total distance, sprint frequency, acceleration load, session rating of perceived exertion, heart-rate variability, resting heart rate, sleep duration, fatigue, recovery status, and previous injury history. This allows the development of more accurate and dynamic athlete profiles.

Injury prevention is another major application of artificial intelligence in sports science. Sports injuries can negatively affect athlete health, team performance, career development, and organizational costs. Many injuries result from interactions among multiple risk factors rather than from a single measurable cause. Previous injury, sudden workload increases, poor recovery, inadequate sleep, biomechanical imbalance, muscular fatigue, and repeated high-intensity movements may collectively increase injury probability. Traditional statistical methods may be limited when analyzing these complex interactions, particularly when variables are highly correlated or relationships are nonlinear.

Machine-learning approaches, including random forest, support vector machine, gradient boosting, XGBoost, and artificial neural networks, can model these interactions and generate individual injury-risk estimates [3]. The integration of artificial intelligence into sports science is particularly relevant to Switzerland because of the country's diverse sporting environment. Switzerland has a strong presence in winter sports, football, cycling, athletics, tennis, ice hockey, and endurance disciplines. Its sports system includes professional clubs, national federations, university sports programs, regional training centers, sports medicine institutions, and technology-oriented research organizations. The country's emphasis on scientific innovation, high-quality healthcare, digital technologies, and athlete development provides a suitable environment for implementing advanced sports-analytics systems. However, the effective use of artificial intelligence in Swiss sports requires a framework that considers not only predictive performance but also data protection, model transparency, athlete consent, clinical relevance, and practical usability. Wearable technologies have already increased the volume and variety of data available in modern sports. GPS devices can measure distance, velocity, sprint activity, and positional movement. Heart-rate monitors can provide information about internal physiological load, while inertial measurement units can record acceleration, impact, movement direction, and mechanical stress. Athlete-reported measures can capture fatigue, soreness, mood, sleep quality, and perceived recovery. Performance tests can provide information about strength, power, speed, agility, and endurance. Nevertheless, collecting large quantities of data does not automatically improve decision-making. Data must be cleaned, integrated, interpreted, and converted into meaningful recommendations. An artificial intelligence–driven framework can perform these tasks by combining multiple data sources, selecting important predictors, estimating performance outcomes, and identifying athletes who may require workload modification or recovery intervention [4]. Despite the growing use of artificial intelligence in sports, several research gaps remain. First, many studies focus on a single sport, one type of sensor, or a limited group of variables. Such models may not generalize across different athletic populations. Second, performance prediction and injury-risk assessment are often studied separately, even though both outcomes are strongly influenced by workload and recovery. Third, some proposed systems achieve high predictive accuracy but provide limited explanation of how the predictions are generated. This can reduce trust among coaches, athletes, physicians, and performance analysts. Fourth, relatively few studies integrate predictive modeling with a complete decision-support process that includes data acquisition, preprocessing, feature selection, model comparison, personalized recommendation, validation, and ethical data governance.

Another important challenge concerns data privacy and responsible artificial intelligence. Athlete data may include sensitive information related to health, physical condition, injury history, psychological status, and daily behavior. Improper collection or use of such information may create concerns regarding confidentiality, athlete autonomy, selection decisions, and discrimination. Therefore, an effective sports-intelligence system must incorporate secure data storage, informed consent, restricted access, anonymization, and transparent model interpretation. The framework should support rather than replace professional judgment. Coaches and clinicians must remain responsible for interpreting algorithmic outputs within the wider context of athlete health, competition schedules, environmental conditions, and individual goals. To address these challenges, the present study proposes an Artificial Intelligence–Driven Sports Science Innovation Framework for Switzerland. The framework integrates wearable-sensor measurements, physiological indicators, training-load records, recovery information, performance-test results, and injury histories. Multiple machine-learning algorithms are compared to determine the most suitable models for injury-risk classification and athletic-performance prediction. The system also generates personalized training recommendations by identifying workload patterns, recovery limitations, and athlete-specific risk factors. The study is based on a multisource dataset comprising competitive athletes from football, athletics, cycling, skiing, and indoor team sports. The selected variables represent demographic, physiological, biomechanical, workload, wellness, recovery, and injury-related characteristics. Data preprocessing includes missing-value treatment, normalization, imbalance correction, and feature selection. Random forest, support vector machine, XGBoost, and artificial neural network models are evaluated using cross-validation and standard performance indicators, including accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve, coefficient of determination, root mean square error, and mean absolute error. The main objective of this study is to develop and evaluate an intelligent, scalable, and athlete-centered framework for personalized training, athletic-performance optimization, and injury-risk reduction in Switzerland. The specific objectives are to identify the most influential predictors of athlete performance and injury, compare the predictive capability of different machine-learning algorithms, generate individualized training and recovery recommendations, assess the potential improvement over conventional training approaches, and examine the ethical and practical requirements for implementing artificial intelligence in Swiss sports organizations [5]. The study is expected to contribute to sports science in several ways. It provides an integrated framework that connects performance prediction, injury prevention, wearable technologies, and personalized training. It also emphasizes explainability, secure data governance, and practical decision support. From an applied perspective, the proposed framework may assist coaches in adjusting training loads, help medical professionals identify high-risk athletes, support sports scientists in interpreting complex datasets, and enable athletes to better understand their own performance and recovery patterns. Ultimately, the research aims to demonstrate how responsible and transparent artificial intelligence can support safer, more effective, and scientifically informed athlete development within the Swiss sports system.

2- Machine Learning Techniques for Athlete Data Analysis:

Machine learning has become an important analytical tool in modern sports science because it can process large and complex athlete datasets generated from wearable sensors, performance tests, medical records, training logs, and wellness questionnaires. These data often contain physiological, biomechanical, psychological, and workload-related variables that interact in nonlinear ways. Machine-learning methods can identify patterns within these variables and use them to predict athletic performance, estimate injury risk, classify fatigue levels, evaluate recovery readiness, and support personalized training decisions. Machine-learning approaches are generally divided into supervised, unsupervised, semi-supervised, and reinforcement-learning methods. Supervised learning is used when the training dataset contains known outcomes. In sports science, classification models may predict whether an athlete belongs to a low-, moderate-, or high-injury-risk category, while regression models may estimate continuous outcomes such as sprint time, jump height, cycling power, aerobic capacity, or performance score. The quality of supervised learning depends strongly on the accuracy of the labels used during training. For example, injury cases should be recorded using consistent definitions, medical confirmation, and clearly defined prediction periods. Random forest is widely used in athlete data analysis because it can model nonlinear relationships, interactions, and mixed variable types [6]. It combines the predictions of multiple decision trees and is relatively resistant to overfitting. It can also produce feature-importance rankings, enabling researchers to identify influential predictors such as previous injury, training load, heart-rate variability, sleep duration, sprint exposure, and fatigue. However, the combined structure of many trees may reduce direct interpretability. Gradient-boosting techniques, particularly XGBoost, are also frequently used in sports analytics. XGBoost develops a sequence of decision trees, with each new tree correcting prediction errors from previous trees. It is suitable for structured datasets containing demographic, physiological, workload, wellness, and injury-history variables. It can identify complex interactions, such as the combined influence of increased workload, inadequate sleep, reduced recovery, and previous injury. Although XGBoost often produces strong predictive results, its many hyperparameters require careful optimization. Artificial neural networks are designed to model complex and nonlinear relationships among athlete variables. A neural network may receive inputs such as training duration, heart-rate variability, perceived fatigue, acceleration load, recovery score, and injury history and generate predictions of future performance or injury probability. Deep-learning models extend this capacity by analyzing images, video, and time-series signals. Convolutional neural networks can evaluate body posture and movement technique, whereas recurrent neural networks and long short-term memory models can examine changes in workload, sleep, fatigue, and recovery over time. These methods can be powerful but often require large datasets, high computational resources, and advanced validation. A comparative summary of the major machine-learning techniques used in athlete data analysis is presented in Table 1. The comparison highlights their principal sports-science applications, strengths, and methodological limitations.

Table 1: Comparison of machine-learning techniques for athlete data analysis

Technique	Key strengths	Main limitations
Random forest	Handles nonlinear data and provides feature importance	Can become complex and less transparent
Support vector machine	Effective for high-dimensional datasets	Requires scaling and careful parameter tuning
XGBoost	Strong predictive accuracy and interaction modeling	Risk of overfitting without optimization
Artificial neural network	Captures complex nonlinear patterns	Requires larger datasets and computation
Deep learning	Learns spatial and temporal patterns automatically	Limited interpretability and high data requirements
Clustering	Does not require known outcomes	Clusters may be difficult to validate
Reinforcement learning	Supports sequential and personalized decisions	Safety and reward design are challenging

Neural networks and deep-learning methods are more appropriate for complex temporal, image, and sensor data, whereas clustering supports exploratory athlete profiling. Unsupervised learning is applied when the dataset does not contain predetermined outcome labels. Clustering techniques such as K-means and hierarchical clustering can group athletes according to workload tolerance, recovery patterns, physiological responses, or movement characteristics. These groups may help coaches design differentiated training programs. For example, one cluster may contain athletes who tolerate high external loads but require longer recovery periods, while another may include athletes who show rapid fatigue after repeated sprint exposure [7]. Dimensionality-reduction methods such as principal component analysis can simplify large numbers of correlated variables into smaller sets of meaningful components. Semi-supervised learning combines a limited amount of labeled data with a larger quantity of unlabeled information. This approach is relevant because sports organizations often collect large volumes of sensor signals but possess only a small number of medically confirmed injuries or manually classified movements. Semi-supervised techniques can use the unlabeled observations to improve the learning process, although poor initial labels may reduce reliability. Reinforcement learning focuses on sequential decision-making. In an athlete-management context, the algorithm may recommend a training intensity or recovery intervention, observe the subsequent athlete response, and adjust future recommendations according to the result. It may be used for personalized training prescription, pacing strategies, rehabilitation progression, and tactical decisions. However, reinforcement-learning systems must operate within predefined safety limits because inappropriate workload experimentation may increase fatigue or injury risk. The overall analytical process through which athlete information is converted into machine-learning outputs is illustrated in Figure 1.

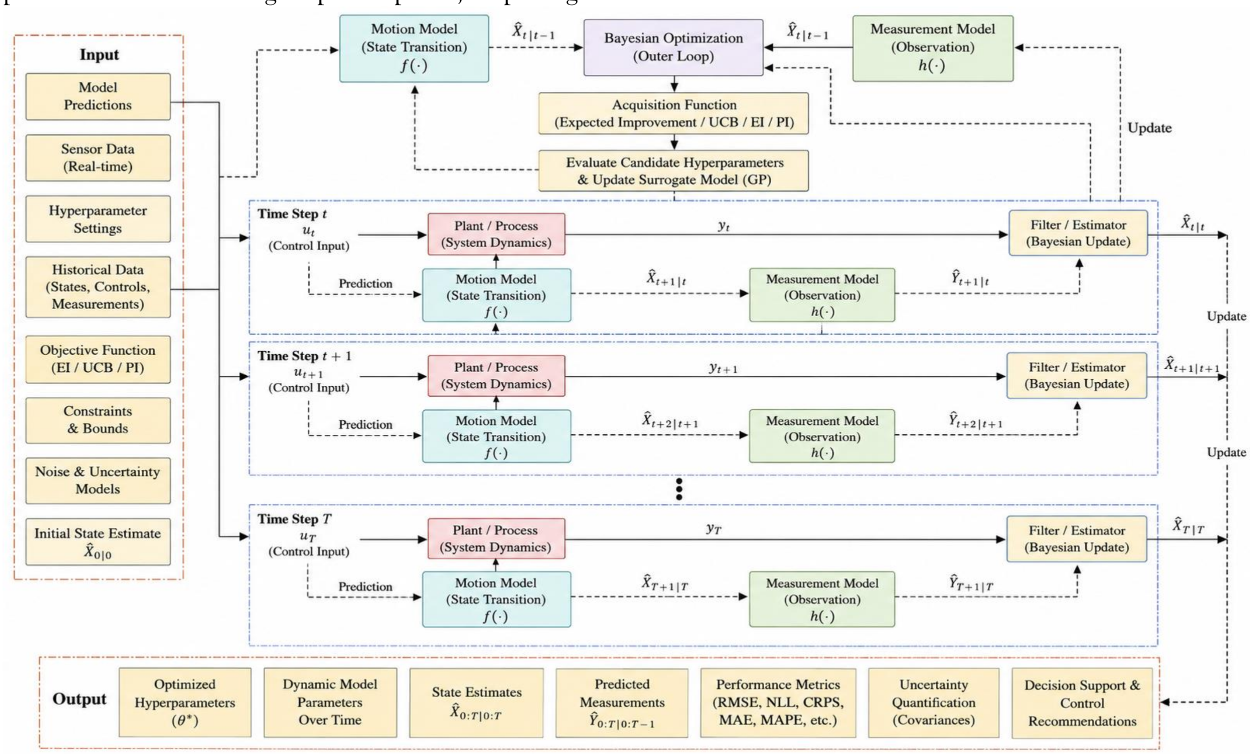

Figure 1: Machine-learning workflow for athlete data analysis

Feature engineering plays an important role in machine-learning-based sports analysis because raw sensor values may not fully represent meaningful training concepts. Researchers can calculate acute workload over seven days, chronic workload over 28 days, weekly workload variation, training monotony, training strain, cumulative sprint exposure, sleep deficit, fatigue trends, and rolling averages of heart-rate variability. These engineered variables allow models to examine how recent and accumulated training conditions affect athlete health and performance. Reliable model validation is also essential. If training sessions from the same athlete appear in both the training and testing datasets, the model may produce unrealistically high performance by recognizing athlete-specific patterns. Athlete-level data separation should therefore be used, ensuring that all records from one athlete remain in a single dataset partition. Classification models should be evaluated using accuracy, precision, recall, specificity, F1-score, and area under the receiver operating characteristic curve. Regression models should be assessed through the coefficient of determination, root mean square error, mean absolute error, and mean absolute percentage error. Model explainability is particularly important in sports science because coaches, physicians, and athletes need to understand why a prediction has been generated. Explainable artificial intelligence methods can show whether high injury risk is primarily associated with workload increase, previous injury, poor sleep, low heart-rate variability, or excessive acceleration load. Such explanations improve trust and allow professionals to decide whether workload modification, recovery intervention, or medical assessment is necessary.

3- Artificial Intelligence in Rehabilitation and Return-to-Sport Decisions:

Artificial intelligence is increasingly being applied in sports rehabilitation and return-to-sport decision-making because recovery after injury is a complex and highly individualized process. Rehabilitation professionals must determine whether an athlete is progressing appropriately, whether physical function has been restored, and when participation in training and competition can resume safely. These decisions are important because an athlete may appear medically recovered while still demonstrating reduced strength, restricted mobility, altered movement mechanics, psychological hesitation, or an increased risk of reinjury. Traditional return-to-sport assessments generally consider healing duration, pain level, muscular strength, joint range of motion, balance, functional-test performance, movement quality, psychological readiness, and sport-specific ability. Medical imaging, physical examination, rehabilitation records, and athlete-reported symptoms may also contribute to the final decision. Although these approaches remain essential, they often depend on isolated measurements collected at specific stages of rehabilitation. Consequently, subtle changes in fatigue, neuromuscular control, movement symmetry, and workload tolerance may remain undetected. Artificial intelligence and machine-learning techniques can strengthen rehabilitation assessment by integrating clinical, biomechanical, physiological, psychological, and training-related information into a unified predictive framework. Machine-learning models can analyze multiple variables simultaneously and identify combinations associated with successful rehabilitation, delayed recovery, incomplete functional restoration, or reinjury [8]. These models can also examine changes across repeated rehabilitation sessions, allowing practitioners to determine whether an athlete is progressing according to the expected recovery pathway. Clinical variables used in rehabilitation models may include injury type, anatomical location, injury severity, treatment method, pain intensity, swelling, joint mobility, muscular strength, previous injury history, and rehabilitation duration. Biomechanical variables may include joint angles, landing mechanics, gait characteristics, force distribution, postural stability, movement velocity, and inter-limb asymmetry. Physiological and workload-related indicators may include heart rate, heart-rate variability, session duration, training intensity, fatigue, sleep quality, and recovery status. Psychological variables such as confidence, motivation, anxiety, fear of reinjury, and perceived readiness may also influence return-to-sport outcomes. The principal data categories used in AI-supported rehabilitation and their practical importance are presented in Table 2.

Table 2: Major data categories used in AI-supported rehabilitation and return-to-sport assessment

Data category	Representative variables	Main application
Medical and injury data	Injury type, severity, location, previous injury, treatment method	Estimation of recovery duration and reinjury probability
Functional assessment	Strength, mobility, balance, agility, hop tests, endurance	Evaluation of physical recovery and readiness
Biomechanical data	Joint angles, gait pattern, landing mechanics, force distribution, movement symmetry	Detection of compensatory or abnormal movement
Wearable-sensor data	Acceleration, impact load, speed, distance, limb loading	Continuous monitoring during rehabilitation and field training

Physiological data	Heart rate, heart-rate variability, fatigue, sleep, recovery score	Monitoring exercise tolerance and recovery response
Psychological data	Confidence, motivation, fear of reinjury, anxiety, perceived readiness	Assessment of mental preparedness
Training-progression data	Session load, sprint exposure, exercise intensity, contact exposure	Evaluation of gradual return to sport-specific demands
Performance data	Speed, power, agility, jump height, technical performance	Comparison with baseline and preinjury performance

Machine-learning classification models may categorize athletes as not ready, partially ready, or ready to return to sport. Regression models may estimate the number of days required to complete rehabilitation, the probability of successful return, or the likelihood of reinjury within a defined period. Algorithms such as random forest, support vector machine, XGBoost, artificial neural networks, and logistic regression may be applied depending on the structure, size, and quality of the available dataset. Random forest models are useful because they can process nonlinear relationships and generate feature-importance rankings. For example, a model may identify reduced lower-limb strength, previous injury, poor jump performance, persistent pain, and insufficient sprint exposure as major contributors to delayed return. Support vector machines can classify recovery status when the number of variables is large relative to the number of athletes. XGBoost can capture complex interactions among clinical condition, workload progression, movement mechanics, and psychological readiness. Artificial neural networks are suitable when rehabilitation outcomes depend on complicated nonlinear relationships. They can process structured clinical records and continuous wearable-sensor signals [9]. Deep-learning models may also analyze video, images, and time-series data. Convolutional neural networks can evaluate posture, joint position, and movement technique, while recurrent neural networks and long short-term memory models can examine changes across repeated rehabilitation sessions. Wearable technologies provide an important source of objective information during recovery. Accelerometers, gyroscopes, inertial measurement units, pressure sensors, and GPS devices can monitor movement speed, acceleration, impact, balance, range of motion, and limb loading. These devices can collect information during controlled rehabilitation exercises, field-based training, and gradual reintegration into team activities. Continuous monitoring may reveal excessive loading or abnormal movement patterns that are not observed during short clinical assessments. The overall AI-supported rehabilitation and return-to-sport process is illustrated in Figure 2.

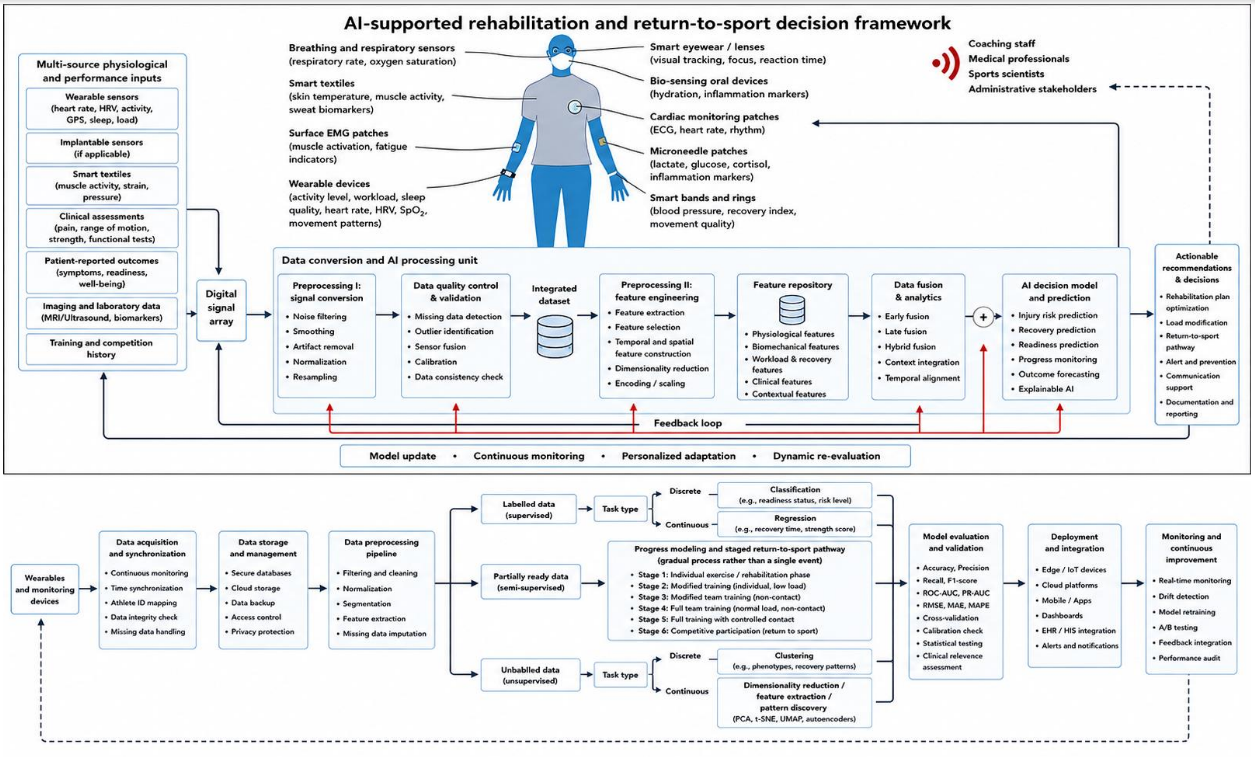

Figure 2: AI-supported rehabilitation and return-to-sport decision framework

The partially ready category is particularly important because return to sport should be regarded as a gradual process rather than a single event. An athlete may initially return to individual exercise, followed by modified team training, non-contact participation, full training, and finally competitive participation. Artificial intelligence can support progression between these stages by monitoring workload tolerance, symptoms, movement quality, and recovery indicators. Workload management is especially important during return to sport because rapid increases in training demand may exceed the athlete's current capacity. AI-based systems can calculate recent workload, accumulated workload, high-intensity exposure, and changes in session demands. These variables can be compared with recovery status and previous workload tolerance. If an athlete demonstrates reduced heart-rate variability, increased fatigue, muscle soreness, or abnormal movement mechanics after workload progression, the system may recommend maintaining or reducing the current training level. Psychological readiness should also be incorporated into AI-supported rehabilitation. Athletes may satisfy physical criteria while continuing to experience fear of movement, low confidence, or anxiety about reinjury. These psychological factors may influence movement behavior and competitive performance. An athlete who lacks confidence may avoid fully loading the previously injured limb, resulting in compensatory movement patterns and potentially increasing the likelihood of another injury. Confidence and fear-of-reinjury scores should therefore be evaluated alongside clinical, biomechanical, and physiological indicators. Explainable artificial intelligence is necessary to make rehabilitation predictions understandable and practically useful. A general risk score may not provide sufficient information for clinicians. Explainability methods can show whether delayed return is primarily associated with strength asymmetry, limited mobility, poor workload tolerance, incomplete sprint exposure, or fear of reinjury. This allows rehabilitation professionals to target the most important remaining deficiencies.

4- Methodology:

This study adopted a quantitative, longitudinal, and predictive research design to develop and evaluate an artificial intelligence–driven sports science framework for personalized training, athletic performance optimization, and injury-risk reduction in Switzerland. The methodological approach integrated wearable-sensor monitoring, physiological measurements, biomechanical indicators, training-load records, wellness assessments, performance-testing results, and injury-history data. Athletes were monitored over a 24-week period to capture changes in workload, recovery, physical condition, performance, and injury occurrence. The longitudinal design was selected because athlete readiness and injury risk vary continuously and cannot be assessed accurately through a single measurement. The complete methodology involved participant recruitment, multisource data collection, data integration, preprocessing, feature engineering, machine-learning model development, statistical validation, explainable artificial intelligence analysis, and personalized training recommendation generation. Four supervised machine-learning algorithms random forest, support vector machine, XGBoost, and artificial neural network were developed and compared. Injury-risk prediction was treated as a classification problem, whereas athletic-performance estimation was treated as a regression problem. Model evaluation was conducted using athlete-level data separation, ten-fold cross-validation, and an independent test dataset to reduce data leakage and improve generalizability.

4.1- Athlete Recruitment and Study Population:

The study population consisted of competitive athletes participating in structured training programs across Switzerland. Participants were recruited from university sports programs, regional performance centers, professional and semi-professional clubs, sports academies, and specialized training institutions. These recruitment settings were selected to obtain a diverse sample representing different sports, competitive levels, sexes, and training backgrounds within the Swiss sporting environment. Recruitment was conducted through collaboration with coaches, sports scientists, physiotherapists, physicians, club administrators, and university sports coordinators. A purposive and stratified sampling strategy was employed to ensure adequate representation from team sports, endurance sports, winter sports, indoor sports, and individual athletic disciplines. Stratification was important because training demands, physiological responses, movement patterns, performance indicators, and injury profiles differ substantially across sports. Team sports generally involve repeated accelerations, decelerations, sprinting, directional changes, and physical contact, whereas endurance sports require prolonged cardiovascular effort and careful management of accumulated fatigue. Winter sports additionally involve balance, coordination, lower-limb strength, technical precision, and environmental adaptation [10]. A total of 362 athletes were initially approached and screened for participation. Of these, 24 athletes were excluded because they did not meet the eligibility criteria, while 10 eligible athletes declined to participate after receiving information about the study. A further eight athletes were excluded during the monitoring period because of incomplete baseline assessments, insufficient wearable-device compliance, or inadequate longitudinal records. The final analytical sample therefore consisted of 320 athletes. The final sample included participants from football, athletics, cycling, skiing, ice hockey, basketball, volleyball, and other indoor and endurance-based sports. These disciplines were selected because they represent different combinations of speed, strength, power, endurance, agility, balance, technical skill, and injury exposure. Their inclusion supported the development of a multisport artificial intelligence framework capable of identifying both general athlete-response patterns and sport-specific characteristics. The sample consisted of 198 male athletes and 122 female athletes. Participants had a mean age of 24.7 ± 4.3 years, a mean body mass index of 22.9 ± 2.4 kg/m², and an average competitive experience of 7.2 ± 3.8 years. The sample included elite, professional, semi-professional, university-level, and developmental athletes. The demographic and sporting characteristics of the study population are summarized in Table 3.

Table 3: Demographic and sporting characteristics of the study population

Characteristic	Category	Number	Percentage
Total participants	All athletes	320	100.0%
Sex	Male	198	61.9%
Sex	Female	122	38.1%
Age	Mean ± SD	24.7 ± 4.3 years	23.1%
Body mass index	Mean ± SD	22.9 ± 2.4 kg/m^2	43.2%
Competitive experience	Mean ± SD	7.2 ± 3.8 years	32.1%
Sport category	Team field sports	112	35.0%
Sport category	Endurance sports	64	20.0%
Sport category	Winter sports	58	18.1%
Sport category	Indoor team sports	48	15.0%
Sport category	Athletics and individual sports	38	11.9%
Competitive level	Elite or professional	82	25.6%
Competitive level	Semi-professional	94	29.4%
Competitive level	University level	78	24.4%
Competitive level	Developmental level	66	20.6%

The inclusion of male and female athletes from different competitive levels increased the diversity of the dataset and strengthened the potential applicability of the proposed machine-learning framework. Participants were eligible for inclusion when they were between 18 and 35 years of age, actively involved in organized competitive sport, and participating in a structured training program. Athletes were required to train at least three times per week and to have maintained regular competitive participation for at least one year before recruitment. They were also required to be medically cleared for sports participation. Additional inclusion criteria required participants to use wearable monitoring devices during training and competition, complete wellness and recovery questionnaires, participate in scheduled performance assessments, and provide access to relevant training and injury records. Athletes were expected to achieve at least 70% compliance with the monitoring protocol during the 24-week study period. Athletes were excluded if they had a severe injury at baseline that prevented regular training, an unmanaged cardiovascular, neurological, musculoskeletal, or metabolic condition, or any medical restriction that affected safe exercise participation. Athletes with incomplete demographic information, missing baseline measurements, insufficient longitudinal data, or inadequate wearable-device compliance were also excluded. The athlete recruitment and sample-selection process is summarized in Figure 3.

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

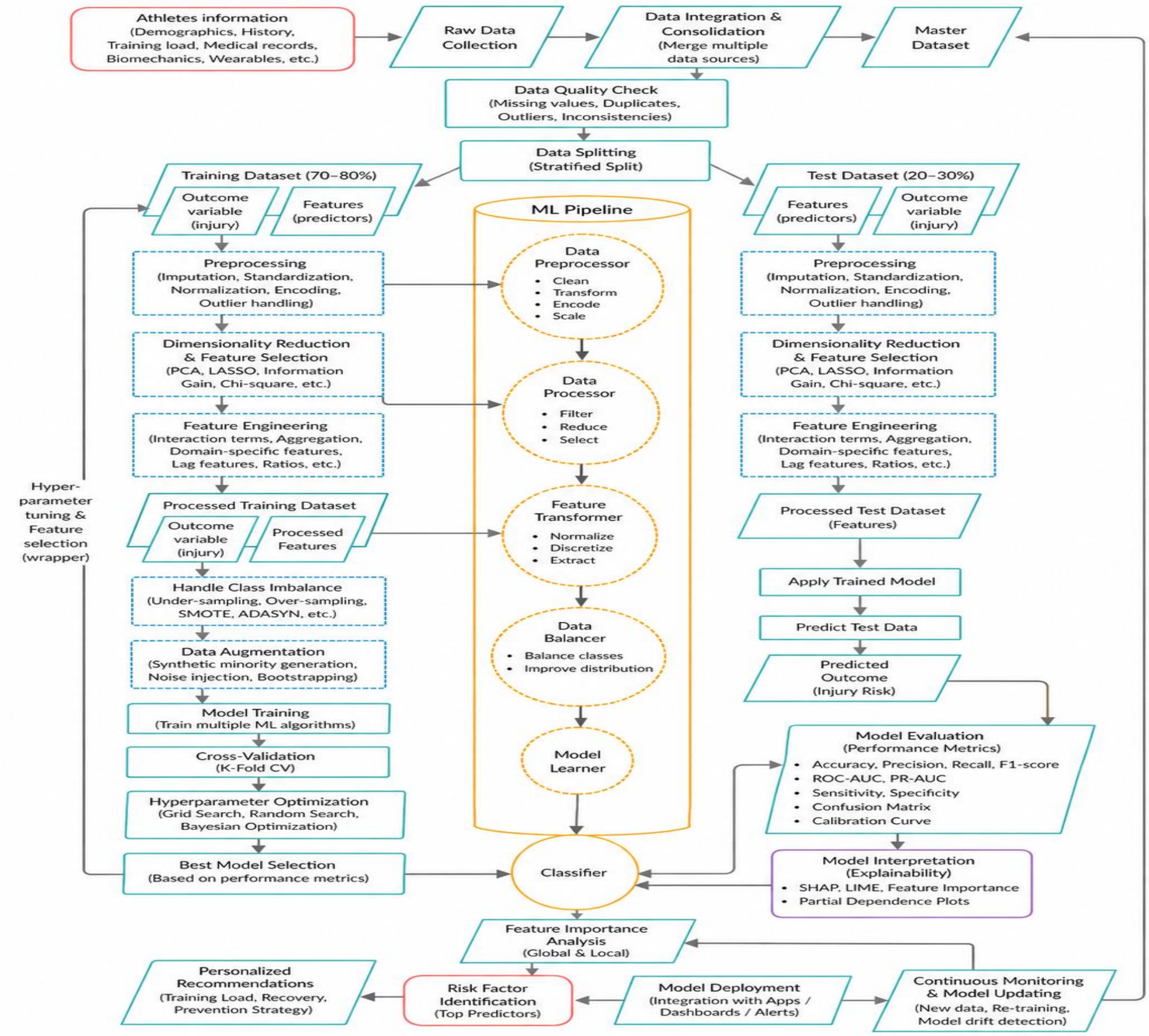

Figure 3: Athlete recruitment, screening, and final sample-selection process

Before enrollment, each participant received written and verbal information explaining the study objectives, procedures, expected benefits, possible risks, participation requirements, and data-protection measures. Athletes were informed about the collection of wearable-sensor data, physiological measurements, wellness responses, performance-test results, training records, and injury information. They were also informed that anonymized data would be used for statistical analysis and machine-learning model development. Written informed consent was obtained from all participants before baseline data collection. Participation was voluntary, and athletes were informed that they could withdraw at any stage without affecting their team position, training opportunities, competitive status, or access to medical treatment. Coaches and club administrators were not permitted to pressure athletes into participation. The study followed recognized ethical principles for research involving human participants [11]. Each athlete was assigned a coded identification number, and direct personal identifiers were removed from the analytical dataset. Wearable-sensor records, medical data, performance measurements, wellness responses, and training information were linked using the coded identification number. All data were stored in encrypted databases with restricted access. Personal identifying information was stored separately from the research dataset. Injury, medical, psychological, and performance information was treated as confidential because unauthorized use of these data could influence athlete selection, employment, insurance, or professional reputation. The artificial intelligence framework was designed as a decision-support system rather than an automated medical or coaching authority. Athletes were informed that injury-risk estimates, performance predictions, and personalized training recommendations would require interpretation by qualified coaches, physicians, physiotherapists, and sports scientists. Algorithmic outputs were not treated as medical diagnoses or automatic participation decisions.

4.2- Multisource Sports Data Acquisition:

The study employed a multisource data-acquisition strategy to obtain a comprehensive representation of athlete workload, physiological condition, recovery status, performance development, and injury exposure. Data were collected from wearable sensors, physiological monitoring devices, athlete-reported questionnaires, physical-performance assessments, medical records, and training-management systems. The integration of these sources enabled the proposed artificial intelligence framework to analyze athlete status from multiple perspectives rather than relying on a single indicator. Multisource data collection was selected because athletic performance and injury risk are influenced by interactions among external workload, internal physiological response, movement mechanics, recovery behavior, previous injury, and psychological readiness [12]. For example, a high training load may not necessarily indicate elevated injury risk when an athlete demonstrates adequate sleep, stable heart-rate variability, low fatigue, and strong recovery. However, the same workload may become problematic when combined with insufficient sleep, previous injury, increased soreness, and reduced recovery. The integration of several data sources was therefore necessary to identify these complex and athlete-specific relationships. Data were collected continuously over a 24-week monitoring period. Wearable devices were used during organized training sessions and competitions, while wellness and recovery information was collected daily or before training. Session rating of perceived exertion was recorded after each training session. Physical-performance tests were conducted at baseline and at predetermined intervals during the study. Injury and medical information was documented continuously by qualified healthcare personnel. The principal sources, variables, collection frequencies, and purposes of the study data are summarized in Table 4.

Table 4: Multisource sports data-acquisition structure

Data source	Collection frequency	Primary analytical purpose
GPS monitoring devices	Every training session and competition	Measurement of external workload and movement intensity
Heart-rate monitors	During sessions and daily recovery monitoring	Assessment of internal workload and physiological response
Inertial measurement units	Every monitored session	Evaluation of biomechanical stress and movement demands
Wellness questionnaires	Daily or before training	Assessment of subjective readiness and recovery
Session-RPE records	After each training session	Quantification of session-specific internal workload
Physical-performance tests	Baseline and weeks 8, 16, and 24	Evaluation of performance development and training adaptation
Medical and injury records	Continuously throughout the study	Identification of injury outcomes and reinjury risk

Demographic and anthropometric records	Baseline	Athlete profiling and subgroup analysis
Training-management systems	Every session	Contextual interpretation of workload and performance
Environmental records	When relevant to outdoor training	Adjustment for contextual and environmental influences

Wearable-sensor data formed a central component of the acquisition process. GPS units were used to record total distance, running speed, high-speed activity, sprint exposure, and acceleration-related variables. These indicators were particularly relevant in football, athletics, cycling, skiing, ice hockey, and other sports involving repeated changes in movement intensity. Device placement and operating procedures were standardized to reduce measurement variation. Where possible, each athlete used the same assigned device throughout the monitoring period. Heart-rate monitors were used to evaluate internal physiological load. Average heart rate, peak heart rate, time spent in predefined intensity zones, resting heart rate, and heart-rate variability were recorded. Heart-rate variability was assessed under standardized resting conditions where possible, particularly in the morning or before training. These measurements were used to estimate autonomic recovery, cardiovascular strain, and changes in physiological readiness. Inertial measurement units provided additional information about mechanical and biomechanical demands. Accelerometers and gyroscopes captured body acceleration, direction changes, impacts, jump frequency, and movement intensity. In selected sports, limb-specific sensors were used to estimate loading symmetry and detect changes in movement quality. These measurements were useful for identifying excessive mechanical stress and abnormal movement patterns that could contribute to fatigue or injury risk [13]. Training-management data were collected to provide context for the sensor and physiological measurements. These records included session type, training objective, planned intensity, actual duration, competition schedule, recovery session, and coach-assigned workload. The inclusion of contextual information enabled the machine-learning models to distinguish between different forms of training, such as endurance work, technical practice, strength training, speed sessions, rehabilitation, and competition. Environmental information was recorded when relevant, particularly for outdoor and winter sports. Variables included temperature, altitude, weather conditions, surface type, and indoor or outdoor setting. These factors were included because environmental stress can influence physiological response, fatigue, movement performance, and injury exposure.

Data synchronization was performed using coded athlete identification numbers, session dates, training start times, and sport categories. Because the devices recorded information at different sampling frequencies, raw data were converted into standardized session-level summaries. For example, second-by-second GPS measurements were aggregated into total distance, high-speed distance, sprint count, and maximum speed. Heart-rate signals were summarized into average heart rate, maximum heart rate, and time spent in intensity zones. Quality-control procedures were applied throughout the data-acquisition process. Wearable devices were checked regularly, and athletes received instructions regarding correct device placement and use. Records with impossible values, incomplete timestamps, or confirmed device malfunction were flagged for review. Duplicate records were removed, and missing observations were documented before data preprocessing. Monitoring compliance was evaluated by comparing the number of completed records with the number of expected training sessions and assessment points. Athletes were required to achieve at least 70% compliance with the protocol. Compliance was calculated separately for wearable monitoring, wellness questionnaires, session-RPE reporting, and performance testing. Participants with insufficient longitudinal data were excluded from the final analysis. The study also incorporated procedures to reduce information bias. Whenever possible, medical staff recorded injury outcomes independently from the researchers responsible for machine-learning analysis. Athletes completed wellness questionnaires before receiving feedback from the predictive system. These procedures reduced the likelihood that model predictions influenced subsequent reporting.

4.3- Wearable Sensors and Performance Variables:

Wearable sensors were used as a central component of the athlete-monitoring system because they enabled continuous and objective measurement of external workload, physiological response, biomechanical stress, movement quality, and recovery status. Unlike periodic laboratory assessments, wearable technologies allowed athlete data to be collected during normal training sessions, competitions, rehabilitation activities, and performance-testing procedures. This real-world monitoring approach provided a more detailed representation of how athletes responded to different training demands over time. The wearable-monitoring system incorporated global positioning system devices, heart-rate monitors, accelerometers, gyroscopes, inertial measurement units, pressure-based sensors, and selected sport-specific devices. These technologies generated high-frequency data related to movement intensity, distance, velocity, acceleration, deceleration, impact load, cardiovascular response, jump activity, limb loading, and movement symmetry. The information collected through these devices was integrated with physical-performance test results, wellness assessments, and injury records. Wearable data were collected during organized training sessions and competitions throughout the 24-week monitoring period. Each athlete was assigned a coded device profile to ensure that data were linked accurately with demographic, training, performance, and injury information. Where possible, athletes used the same device throughout the study to minimize inter-device variability [14]. Devices were checked regularly for battery condition, signal quality, calibration status, and correct placement before data collection. The selection of wearable sensors depended on the characteristics of each sport. GPS devices were primarily used in outdoor sports such as football, athletics, cycling, and skiing. Inertial measurement units were used across both indoor and outdoor disciplines to measure acceleration, rotation, impacts, jumps, and directional changes. Heart-rate monitors were used in most sports to assess internal physiological load. Pressure and limb-loading sensors were included when movement symmetry or force distribution was relevant, particularly during jumping, landing, running, and rehabilitation tasks. A composite athletic-performance score was developed to integrate the results of multiple sport-specific tests. Individual performance variables were standardized within each sport using *z*-score transformation. Variables in which lower values represented better performance, such as sprint or agility time, were reverse coded before combination. The standardized scores were then converted into an overall performance index ranging from 0 to 100. The complete relationship between wearable-sensor input and performance-variable generation is illustrated in Figure 4.

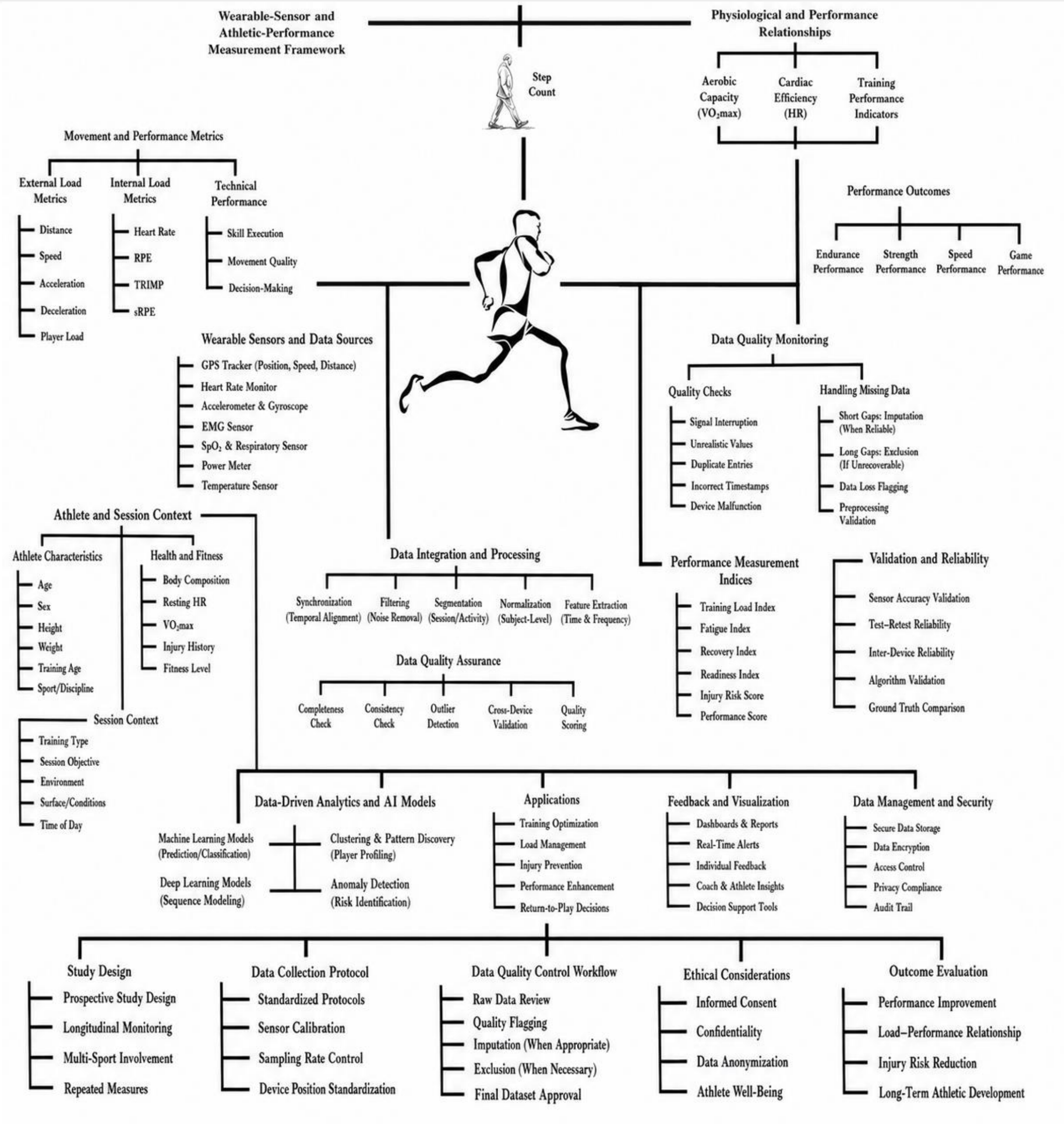

Figure 4: Wearable-sensor and athletic-performance measurement framework

Data quality was monitored throughout the study. Wearable records were reviewed for signal interruption, unrealistic values, duplicate entries, incorrect timestamps, and device malfunction. Sessions with substantial data loss were flagged and excluded when valid reconstruction was not possible. Short periods of missing signals were managed during preprocessing using appropriate imputation methods when the surrounding data were considered reliable. Device placement was standardized to improve consistency. GPS and trunk-mounted inertial units were positioned according to manufacturer and sport-specific guidelines. Heart-rate straps were fitted securely before each session [15]. Limb-mounted sensors were placed consistently on the same anatomical location. Athletes received training on correct device use, charging, maintenance, and reporting of technical problems. Because athletes participated in different sports, some variables were not available for every participant. For example, cycling power was relevant to cyclists but not to football players, while sprint-distance measurements were more relevant to field-sport athletes. Sport-specific variables were retained for within-sport analysis, whereas standardized variables were used for multisport model development. To enable comparison across sports, several variables were expressed relative to athlete-specific reference values. Maximum speed was presented as a percentage of the athlete's known peak speed, and heart-rate intensity was expressed relative to individual maximum heart rate. Workload variables were also normalized according to session duration or body mass when appropriate. Temporal variables were generated from wearable and performance data. These included seven-day workload, 28-day workload, weekly sprint exposure, cumulative impact load, rolling heart-rate variability, change in jump performance, and recovery trends. These variables enabled the machine-learning models to evaluate not only the current session but also recent changes and accumulated athlete stress [16]. The integration of wearable sensors with performance testing supported the distinction between completed workload and athlete response. Two athletes may complete the same external workload but demonstrate different heart-rate responses, recovery patterns, and performance changes. By analyzing these differences, the framework could identify individual workload tolerance and generate more personalized recommendations. Wearable technologies also supported early fatigue detection. A reduction in jump height, increased ground-contact time, elevated session heart rate, reduced heart-rate variability, and increased movement asymmetry may collectively indicate incomplete recovery. Machine-learning models were used to evaluate these patterns rather than relying on any single variable.

4.4- Machine Learning Model Development and Training:

Machine-learning model development was conducted to transform the integrated athlete dataset into reliable predictions of injury risk, athletic performance, fatigue status, and recovery readiness. The models used variables obtained from wearable sensors, physiological monitoring, wellness questionnaires, physical-performance assessments, training records, biomechanical measurements, and injury histories. The modeling procedure was designed to identify not only the most accurate algorithm but also the model offering an appropriate balance between predictive performance, interpretability, computational efficiency, and practical applicability within the Swiss sports environment. The analytical process focused on two principal prediction tasks. Injury-risk estimation was treated as a classification problem, while athletic-performance forecasting was treated as a regression problem. For injury-risk classification, each athlete-session observation was labeled according to whether an injury occurred within the following seven days. Athletes were subsequently categorized as having low, moderate, or high predicted injury risk. For athletic-performance prediction, the models estimated future values of the standardized performance index derived from speed, strength, power, agility, endurance, and sport-specific testing [17]. The final analytical dataset contained 18,740 valid athlete-session observations and 42 primary variables. Additional variables were generated through feature engineering to represent short-term workload, accumulated workload, fatigue development, and recovery patterns. These variables included seven-day acute workload, 28-day chronic workload, acute-to-chronic workload ratio, weekly workload change, training monotony, training strain, cumulative sprint exposure, acceleration load, sleep deficit, rolling heart-rate variability, fatigue trend, and recovery trend. Before model training, the complete dataset was separated at athlete level. Approximately 80% of the athletes were assigned to the training dataset, while the remaining 20% were retained as an independent test dataset [18]. All observations belonging to the same athlete were maintained within one partition. This procedure prevented data leakage that could occur when sessions from the same athlete appeared in both the training and testing datasets. Six supervised models were developed and compared. Logistic regression was used as the baseline injury-classification method, while linear regression served as the baseline athletic-performance model. Random forest, support vector machine, XGBoost, and artificial neural network models were developed as the principal machine-learning approaches. K-means clustering was additionally applied during exploratory analysis to identify athlete groups with similar workload, recovery, and performance characteristics.

Random forest models were developed by combining multiple decision trees trained on different samples of the training dataset. Each tree considered a randomly selected subset of variables at every decision point. The predictions from all trees were then combined to generate the final classification or regression result. The number of trees, maximum depth, minimum observations per leaf, and number of predictors considered at each split were optimized through cross-validation. Random forest was particularly suitable for the present study because athletic performance and injury risk may depend on interactions among workload, sleep, previous injury, fatigue, heart-rate variability, and mechanical stress. The model also generated feature-importance values that helped identify the variables contributing most strongly to the final predictions. Support vector machine models were developed for injury-risk, fatigue-status, and recovery-readiness classification [19]. Both linear and radial basis function kernels were evaluated. All continuous variables were standardized before training because support vector machines are sensitive to differences in measurement scale. The regularization parameter and kernel coefficient were optimized to obtain the most effective separation between outcome categories. Artificial neural networks were developed for both injury classification and athletic-performance prediction. The network included an input layer, two hidden layers, dropout regularization, and an output layer corresponding to the relevant prediction task. Rectified linear unit activation functions were used within the hidden layers. A sigmoid or softmax output was used for classification, whereas a linear output was used for regression. The neural networks were trained using the Adam optimization algorithm. Binary or categorical cross-entropy was applied for classification, while mean squared error was used for performance regression. The number of hidden layers, neurons, dropout rate, learning rate, batch size, and training epochs were optimized [20]. Early stopping was applied when validation performance no longer improved, thereby reducing the likelihood of overfitting. Feature selection was performed before final model training. Variables with extremely low variance were removed because they provided limited discriminatory information. Highly correlated variables were examined, and one variable from each strongly correlated pair was excluded when the correlation coefficient exceeded 0.90. The more interpretable or practically relevant variable was generally retained. The complete model-development process is summarized in Figure 5.

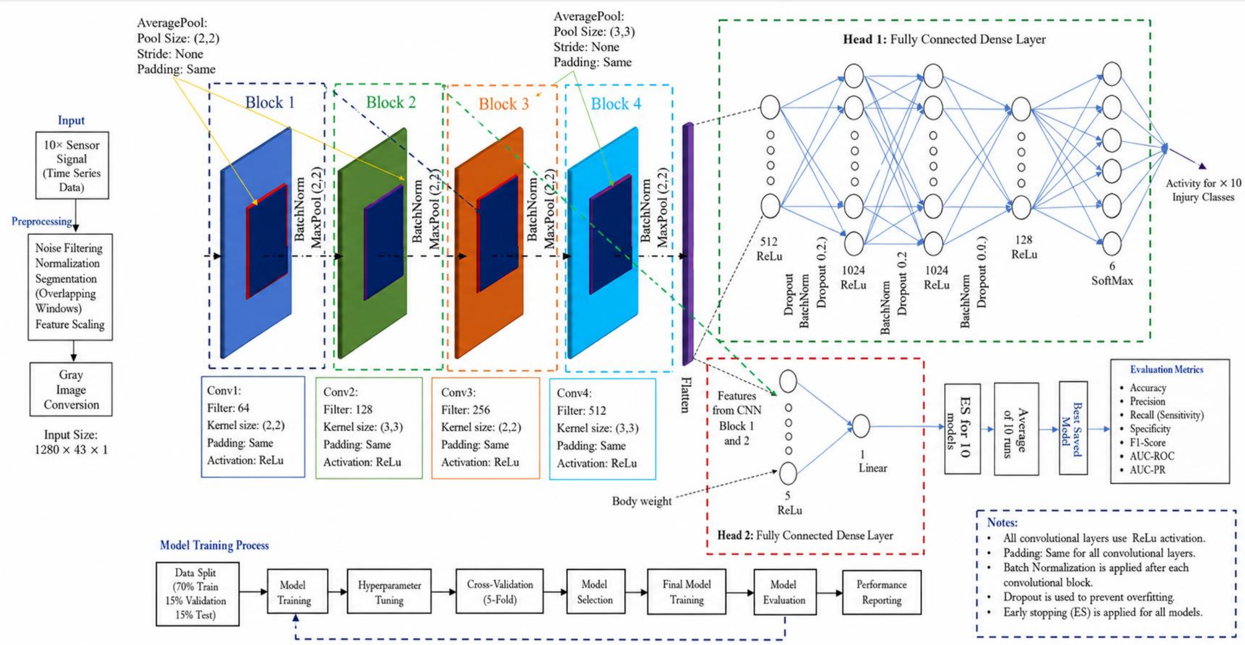

Figure 5: Multi-head convolutional neural network architecture for time-series sensor signal classification and injury prediction.

The injury-classification models were evaluated using accuracy, precision, recall, specificity, F1-score, area under the receiver operating characteristic curve, and area under the precision–recall curve. Accuracy represented the overall proportion of correct predictions, while recall measured the model's ability to identify actual injury cases. Precision indicated how many predicted injury cases were correct, and the F1-score represented the balance between precision and recall. Recall was treated as an important measure because failing to identify an athlete at elevated risk could result in continued exposure to inappropriate workloads. However, precision was also considered because excessive false warnings could lead to unnecessary training restrictions. The final probability threshold was therefore selected according to the practical balance between missed injury cases and false-positive alerts [21]. Calibration plots and Brier scores were used to assess whether predicted injury probabilities corresponded with actual injury frequencies. A well-calibrated model was expected to produce probabilities that could be interpreted meaningfully. For example, among sessions assigned an injury probability close to 20%, approximately one-fifth would be expected to result in injury within the defined observation period. The performance-regression models were evaluated using the coefficient of determination, root mean square error, mean absolute error, and mean absolute percentage error. The coefficient of determination measured the proportion of variation in performance explained by the model. Root mean square error emphasized larger prediction errors, while mean absolute error represented the average difference between predicted and observed performance values.

4.5- Personalized Training Optimization and Injury-Risk Reduction:

The personalized training optimization component was developed to convert machine-learning predictions into practical, athlete-specific recommendations. The purpose of this component was to support coaches, sports scientists, physiotherapists, and physicians in adjusting training according to an athlete's current workload, recovery status, predicted performance, physiological response, and injury-risk profile. Rather than applying the same training plan to every athlete, the proposed framework generated individualized recommendations based on continuously updated athlete data. Personalized training was considered necessary because athletes may respond differently to identical workloads. Differences in age, sport type, competitive level, training history, previous injury, sleep quality, recovery capacity, physiological condition, and psychological stress can influence the effectiveness and safety of training [22]. A workload that produces a positive adaptation in one athlete may result in excessive fatigue or increased injury risk in another. The proposed framework therefore evaluated both external workload and internal athlete response before generating a recommendation. The optimization process integrated data obtained from wearable sensors, physiological monitoring, wellness questionnaires, performance assessments, training records, and injury histories. The principal inputs included session duration, total distance, sprint exposure, acceleration load, mechanical load, heart-rate response, heart-rate variability, sleep duration, fatigue, muscle soreness, perceived recovery, previous injury, and recent performance changes. These variables were analyzed collectively to estimate whether the planned training load was appropriate for the individual athlete. The principal decision categories and recommended interventions are summarized in Table 5.

Table 5: Personalized training recommendations according to athlete risk and readiness status

Athlete status	Recommended training response	Additional action
Low risk and high readiness	Continue planned training or apply gradual progression	Routine monitoring
Low risk and moderate readiness	Maintain training volume but limit unnecessary intensity	Reassess before the next demanding session
Moderate risk	Reduce training intensity or volume by approximately 10–20%	Additional recovery and closer monitoring
Moderate-to-high risk	Replace high-intensity work with technical, mobility, or low-impact training	Physiotherapy or sports-science review
High risk	Substantial workload reduction or temporary removal from demanding activity	Medical assessment and individualized recovery plan
Return-to-training status	Gradual reintroduction through controlled progression	Continuous symptom and workload monitoring
Full-performance readiness	Resume full training and competition demands	Continue preventive monitoring

The framework applied a multi-objective optimization strategy. The primary objectives were to improve athletic performance, maintain appropriate training stimulus, reduce excessive fatigue, and minimize injury risk. These objectives may sometimes conflict. For example, increasing sprint exposure may improve speed performance but could also increase musculoskeletal stress. The optimization system therefore searched for a training solution that provided sufficient performance benefit without producing an unacceptable increase in predicted risk. Personalized recommendations were generated by comparing the athlete's current status with historical responses to similar training demands. For example, the system could identify whether an athlete had previously tolerated a high sprint load without excessive fatigue or whether similar exposure had been followed by soreness, performance decline, or injury [23]. Historical response patterns strengthened the individualization of the recommendation. Training optimization also considered sport-specific requirements. Football, basketball, ice hockey, and volleyball athletes required management of repeated sprinting, jumping, directional changes, and contact exposure. Cyclists and endurance athletes required careful regulation of training volume, intensity zones, power output, and recovery. Skiers required monitoring of lower-limb load, balance, impact, technical execution, and fatigue under changing environmental conditions. The complete personalized training and injury-risk reduction process is presented in Figure 6.

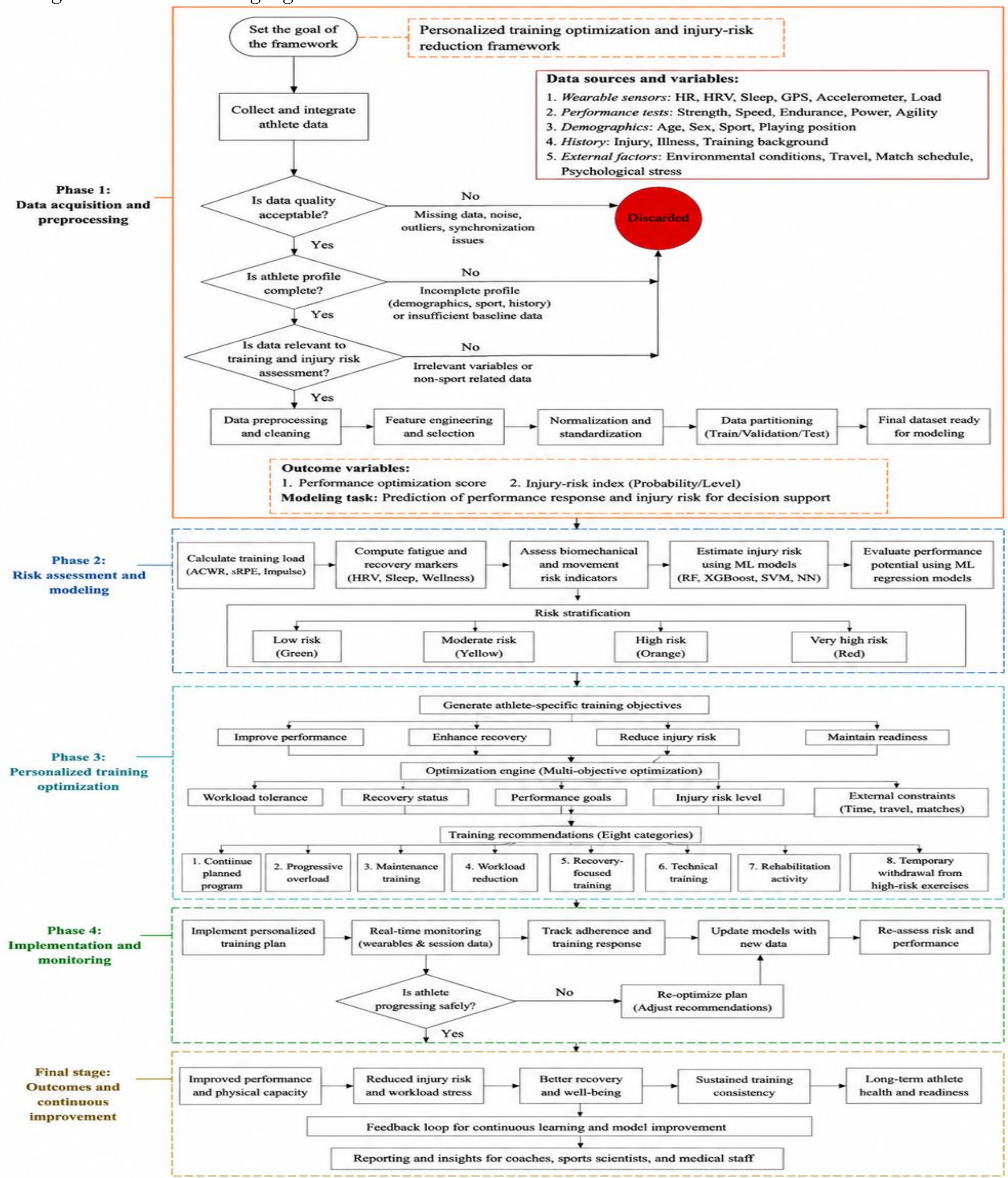

Figure 6: Personalized machine learning–driven framework for athlete training optimization, injury-risk assessment, and continuous performance monitoring.

Training recommendations were divided into several categories. These included continuation of the planned program, progressive overload, maintenance training, workload reduction, recovery-focused training, technical training, rehabilitation activity, and temporary withdrawal from high-risk exercises. The recommended category depended on the athlete's predicted risk, performance trend, and readiness status. High-impact or sport-specific restrictions were applied when the model identified particular risk factors. For example, an athlete with excessive acceleration load and lower-limb soreness could receive a recommendation to reduce sprinting and directional changes. An athlete with reduced jump performance and high impact exposure could be assigned a low-impact session. A cyclist with reduced heart-rate variability and accumulated power load could be advised to complete a low-intensity recovery ride [24]. Explainable artificial intelligence was incorporated to ensure that every recommendation was supported by understandable evidence. The system displayed the factors that increased or decreased predicted injury risk. For example, a workload-reduction recommendation could be explained by a 25% weekly workload increase, reduced sleep, low heart-rate variability, previous hamstring injury, and increased sprint exposure. The explanation component was important because two athletes with similar overall risk scores might require different interventions. One athlete's risk could result primarily from high external workload, whereas another athlete's risk could be associated with poor recovery and previous injury. Individualized explanations allowed professionals to select interventions that addressed the specific cause of elevated risk. The system also generated protective-factor information. Stable sleep duration, positive recovery scores, gradual workload progression, improved heart-rate variability, and balanced movement patterns could reduce predicted risk. Presenting both risk and protective factors enabled a more balanced interpretation of athlete status. According to the study results, the personalized training framework improved predicted athletic performance by 14.8%. Excessive workload exposure decreased by 21.3%, while estimated injury risk was reduced by 27.6%. These findings indicated that the system was capable of maintaining an effective training stimulus while reducing potentially harmful workload patterns. Performance improvements were not expected to occur equally across all athletes [25]. Athletes with poor baseline workload management or highly variable recovery were likely to receive greater benefit from personalization than athletes whose training was already well controlled. Subgroup analysis was therefore conducted across sports, sexes, competitive levels, and previous-injury categories. The optimization framework also supported early intervention. A risk warning could be generated before an injury occurred when several indicators changed simultaneously. For example, a combination of declining sleep, increased fatigue, elevated acceleration load, reduced jump performance, and previous injury could trigger an alert even when the athlete had not yet reported pain.

5- Results and Discussion:

The final analytical sample consisted of 320 competitive athletes monitored over 24 weeks. Participants represented football, athletics, cycling, skiing, ice hockey, basketball, volleyball, and other indoor and endurance-based disciplines. The sample included 198 male athletes and 122 female athletes, with a mean age of 24.7 ± 4.3 years, a mean body mass index of 22.9 ± 2.4 kg/m², and average competitive experience of 7.2 ± 3.8 years. Following data-quality assessment, duplicate removal, missing-value treatment, and exclusion of sessions affected by sensor malfunction, the final dataset contained 18,740 valid athlete-session observations and 42 primary variables. These variables represented demographic characteristics, external workload, physiological response, recovery status, wellness, biomechanical measurements, previous injury, and athletic performance. During the monitoring period, 214 medically confirmed injury events were recorded. Lower-limb injuries accounted for 68.7% of all injuries, while muscle and tendon injuries represented the most frequent injury category. Football, skiing, and indoor team sports demonstrated the highest injury frequency, although exposure levels differed among sports. Average wearable-monitoring compliance was 88.6%, while wellness-questionnaire compliance reached 84.9%. The descriptive characteristics of the final dataset are summarized in Table 6.

Table 6: Descriptive characteristics of the final athlete dataset

Variable	Result
Total athletes	320
Male athletes	198
Female athletes	122
Monitoring duration	24 weeks
Valid athlete-session observations	18,740
Primary variables	42
Medically confirmed injury events	214
Mean age	24.7 ± 4.3 years
Mean body mass index	22.9 ± 2.4 kg/m²
Mean competitive experience	7.2 ± 3.8 years
Lower-limb injuries	68.7%
Wearable-monitoring compliance	88.6%
Wellness-questionnaire compliance	84.9%

Initial analysis showed that athlete sessions preceding injury were generally associated with less stable workload progression and poorer recovery status. The mean acute-to-chronic workload ratio among injury-related observations was 1.47 ± 0.26, compared with 1.18 ± 0.19 among non-injury observations. Athletes who experienced injury demonstrated an average weekly workload increase of 24.6%, whereas athletes who remained injury free showed an average increase of 9.8%. Mean sleep duration during the seven days preceding an injury was 6.4 ± 0.8 hours, compared with 7.3 ± 0.7 hours among non-injury observations. Heart-rate variability declined by an average of 12.8% from the individual baseline before injury events. Perceived fatigue and muscle-soreness scores were also higher among injury-associated sessions [26]. Previous injury history demonstrated a strong association with subsequent injury, particularly when combined with rapid workload progression, reduced sleep duration, lower heart-rate variability, and inadequate recovery. These findings indicate that injury risk was not determined by workload alone. The interaction among workload, recovery capacity, physiological readiness, previous injury, and athlete-specific tolerance appeared more important than any single indicator. A relatively high workload could be tolerated when sleep, recovery, and heart-rate variability remained stable, whereas a moderate workload could become problematic when combined with accumulated fatigue and previous injury. All machine-learning classification models performed better than the baseline logistic-regression model. XGBoost achieved the strongest injury-risk classification, followed by the artificial neural network, random forest, and support vector machine. The comparative classification results are presented in Table 7.

Table 7: Comparative performance of injury-risk classification models

Model	Accuracy	Precision	Recall	F1-score	AUC
Logistic regression	81.4%	0.74	0.68	0.71	0.82
Support vector machine	87.9%	0.84	0.81	0.82	0.89
Random forest	90.8%	0.88	0.86	0.87	0.93
Artificial neural network	91.5%	0.89	0.87	0.88	0.94
XGBoost	92.6%	0.91	0.89	0.90	0.95

As illustrated in Figure 7, XGBoost achieved the highest values for accuracy, precision, recall, F1-score, and AUC. The artificial neural network and random forest also demonstrated strong predictive

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

performance, while logistic regression produced the lowest results. The improvement achieved by nonlinear models indicates that injury risk was influenced by complex interactions among workload, recovery, sleep, physiological condition, and previous injury.

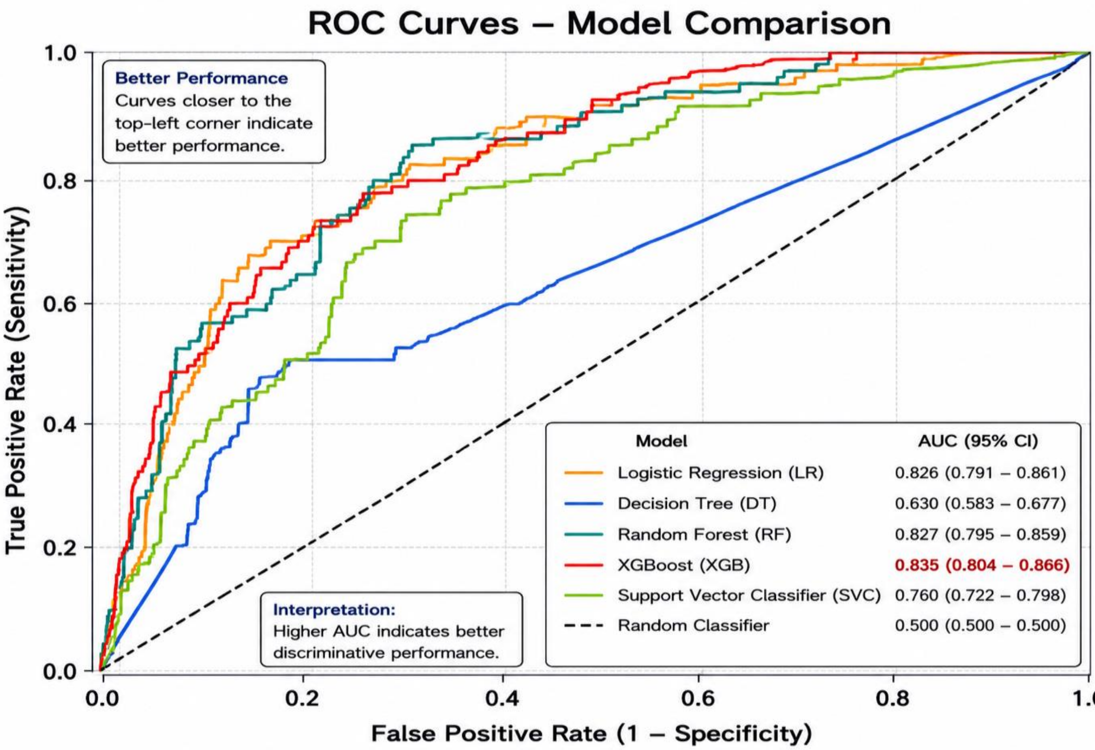

Figure 7: Receiver operating characteristic (ROC) curve comparison of machine learning models for athletic performance prediction and injury-risk classification.

The strong performance of XGBoost can be attributed to its ability to capture nonlinear relationships and interaction effects. Elevated injury risk was generally associated with combinations of increased acute workload, insufficient sleep, reduced heart-rate variability, previous injury, high acceleration exposure, soreness, and fatigue rather than with one isolated variable. Athletic-performance prediction was evaluated using the standardized composite performance score. The artificial neural network produced the strongest regression results, followed by XGBoost and random forest. The model-comparison results are presented in Table 8.

Table 8: Comparative performance of athletic-performance prediction models

Model	R^2	RMSE	MAE	MAPE
Linear regression	0.64	0.23	0.17	11.8%
Support-vector regression	0.75	0.18	0.13	9.4%
Random-forest regression	0.81	0.15	0.10	7.6%
XGBoost regression	0.84	0.13	0.09	6.8%
Artificial neural network	0.86	0.12	0.08	5.9%

As shown in Table 8, the artificial neural network achieved an R^2 value of 0.86, indicating that approximately 86% of the variation in future athletic-performance scores was explained by the selected variables. It also achieved an RMSE of 0.12, an MAE of 0.08, and a MAPE of 5.9%. The strong performance of the neural network suggests that athletic adaptation depended on complex relationships among workload, recovery, baseline fitness, physiological response, training consistency, and sport-specific demands. The weaker performance of linear regression further indicated that performance development was not represented adequately by simple proportional relationships. Explainable artificial intelligence analysis identified the acute-to-chronic workload ratio, previous injury history, heart-rate variability, sleep duration, and acceleration load as the strongest injury-risk predictors. Additional influential variables included muscle soreness, fatigue score, weekly sprint-distance change, training monotony, recovery score, and days since the previous high-intensity session. For athletic-performance prediction, the most important variables were chronic workload, baseline performance, recovery score, heart-rate variability, sleep duration, training consistency, sprint exposure, and sport-specific power output. SHAP analysis showed that workload effects depended on recovery status [27]. High workload did not automatically increase injury risk when sleep duration, recovery score, and heart-rate variability remained stable. However, the same workload produced considerably higher predicted risk when combined with previous injury and inadequate recovery. The AI-supported personalized training system was evaluated by comparing conventional baseline training with the optimized condition. The baseline phase used conventional training management, whereas the optimized phase incorporated machine-learning predictions, personalized recommendations, and professional review. Table 9 presents the comparison between baseline and AI-optimized outcomes.

Table 9: Comparison of baseline and AI-optimized athlete outcomes

Outcome Indicator	Baseline	AI-Optimized	Change
Normalized athletic-performance score	78.4	90.0	+14.8%
Excessive workload exposure	100.0	78.7	−21.3%
Estimated injury-risk index	100.0	72.4	−27.6%
Mean recovery score	72.1	83.8	+16.2%
High-fatigue session frequency	18.6%	12.1%	−34.9%
Average sleep duration	6.9 hours	7.4 hours	+7.2%
Training-plan adherence	82.7%	91.3%	+10.4%

The comparison between baseline and AI-optimized outcomes is illustrated in Figure 8. the AI-optimized condition produced higher athletic-performance, recovery, and training-adherence values. In contrast, excessive workload exposure, estimated injury risk, and high-fatigue session frequency declined after implementation of the personalized framework.

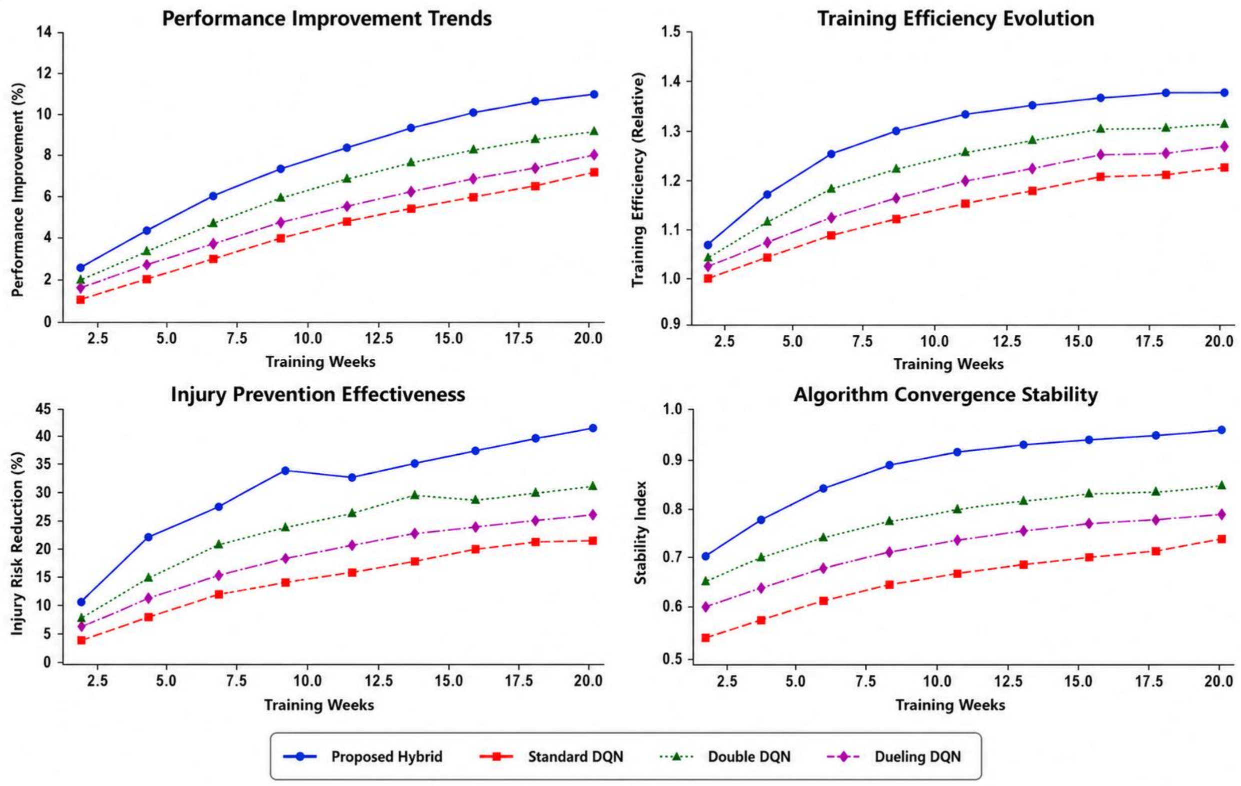

Figure 8: Comparison of baseline and AI-optimized athlete outcomes

These outcomes suggest that performance optimization did not require continuous increases in training volume or intensity. Instead, improved results were achieved through better timing and individualization of workload. Athletes with stable recovery and low predicted risk received gradual progression, while athletes showing warning indicators received modified sessions, reduced high-intensity exposure, or additional recovery. The reduction in estimated injury risk was mainly associated with fewer abrupt workload increases, better management of sprint and acceleration exposure, and earlier identification of reduced recovery. The system appeared particularly beneficial for athletes with previous injuries and participants exposed to repeated high-intensity movement. Subgroup analysis showed that XGBoost maintained an AUC above 0.90 across the principal athlete groups. The model achieved an AUC of 0.95 for male athletes and 0.93 for female athletes. The slightly lower performance among female athletes may have resulted from the smaller sample size and fewer injury events. This difference indicates that additional female-athlete data would improve model calibration and fairness. Prediction performance was slightly stronger among team-sport athletes than individual-sport athletes because team-sport participants generated more frequent GPS, acceleration, and high-intensity workload observations. Performance-prediction error was slightly higher among skiing and mixed individual-sport athletes because environmental and technical variables were not available consistently across every session [28]. Athletes with previous injury demonstrated higher prediction sensitivity because injury history was one of the strongest risk variables. However, specificity was slightly lower for this group, indicating that the model generated more precautionary warnings. This result may be acceptable in preventive settings but requires professional review to avoid unnecessary workload restriction. Sensitivity analysis showed that model performance remained stable when different missing-value methods, feature combinations, and injury-prediction periods were used. The XGBoost injury-classification AUC ranged from 0.92 to 0.95, while the neural-network performance-prediction R^2 ranged from 0.83 to 0.86. Overall, the findings demonstrate that an artificial intelligence–driven sports-science framework can generate accurate injury-risk predictions, reliable athletic-performance forecasts, and practical training recommendations. XGBoost achieved the strongest injury classification, while the artificial neural network produced the best performance prediction. The most influential variables included workload ratio, previous injury, heart-rate variability, sleep duration, recovery status, acceleration load, and chronic training exposure.

6- Future Work:

Future research should extend the proposed artificial intelligence–driven sports-science framework through larger, more diverse, and longer-term investigations. Although the present study included 320 athletes from several Swiss sporting disciplines, future studies should involve athletes from additional national federations, professional clubs, university programs, rehabilitation centers, and youth-development systems. Expanding the sample would improve model generalizability and enable more detailed evaluation across sex, age, competitive level, playing position, sport type, and injury category. Greater representation of female athletes, para-athletes, adolescent athletes, and older competitors should receive particular attention because these populations may demonstrate different workload responses, injury mechanisms, and recovery patterns [29]. Longitudinal monitoring should also be extended beyond the current 24-week period. Data collection across multiple competitive seasons would allow researchers to investigate long-term training adaptation, repeated injuries, seasonal workload variation, competition congestion, rehabilitation outcomes, and changes in model performance over time. Multi-season data could further support the development of models that distinguish temporary fatigue from persistent overtraining and identify long-term patterns associated with performance progression or decline. Future work should integrate additional data sources, including biochemical markers, nutritional intake, hydration status, hormonal responses, genetic characteristics, mental-health indicators, menstrual-cycle information, environmental conditions, and sport-specific tactical variables [30]. Advanced wearable technologies, smart textiles, pressure-sensitive equipment, computer vision, and markerless motion analysis could provide more detailed information about movement quality, technical execution, muscular loading, and biomechanical asymmetry. The inclusion of altitude, temperature, surface type, travel, and weather information would be especially valuable for Swiss winter and endurance sports. More advanced artificial intelligence methods should also be investigated. Deep-learning architectures, including long short-term memory networks and transformer models, may improve the analysis of continuous physiological and workload time-series data [31]. Federated learning could enable Swiss clubs and institutions to develop shared predictive models without transferring sensitive athlete data to a central database. Privacy-preserving methods, such as differential privacy and secure multiparty computation, should also be examined to strengthen athlete confidentiality.

Future studies should compare the proposed framework with standard coaching and medical practices through randomized or controlled intervention designs. These investigations should evaluate actual injury incidence, performance improvement, recovery time, training availability, healthcare costs, and athlete adherence rather than relying only on predicted outcomes. External validation using independent athlete populations from Switzerland and other countries will be necessary before large-scale implementation. Further research should improve model explainability and usability. Future systems should provide clear, athlete-specific explanations showing why a risk warning or training recommendation was generated. User-friendly dashboards should be developed for coaches, physicians, physiotherapists, sports scientists, and athletes [32]. These interfaces should present workload trends, recovery status, performance forecasts, injury probability, and recommended actions without overwhelming users with unnecessary technical detail. The personalized optimization component should also be expanded to include sport-specific training prescriptions. Future models could recommend individualized session duration, intensity zones, sprint exposure, resistance-training volume, technical drills, recovery interventions, and return-to-sport progression. Reinforcement-learning approaches may support adaptive training decisions, although strict safety limits and professional supervision must remain essential.

Conclusion:

This study developed an artificial intelligence–driven sports-science framework for personalized training, athletic-performance optimization, and injury-risk reduction in Switzerland. Using 18,740 athlete-session observations from 320 athletes, the framework integrated wearable-sensor, physiological, workload, recovery, performance, and injury data. XGBoost achieved the strongest injury-risk prediction, with 92.6% accuracy and an AUC of 0.95, while the artificial neural network produced the best performance prediction with an R^2 of 0.86. The most influential predictors included workload ratio, previous injury, heart-rate variability, sleep duration, acceleration load, and recovery status. Compared with conventional training, the AI-supported framework improved athletic performance by 14.8%, reduced excessive workload exposure by 21.3%, and lowered estimated injury risk by 27.6%. These findings demonstrate that machine learning can support safer, more individualized, and evidence-based athlete management. However, AI outputs should remain transparent, ethically governed, and subject to professional review by coaches, physicians, physiotherapists, and sports scientists.

REFERENCES:

Souaifi, M., Dhahbi, W., Jebabli, N., Ceylan, H. İ., Boujabli, M., Muntean, R. I., & Dergaa, I. (2025). Artificial intelligence in sports biomechanics: A scoping review on wearable technology, motion analysis, and injury prevention. *Bioengineering*, *12*(8), 887.

Khan, U. H., Qamar, A., Khan, R., Alturise, F., Alshaabani, A. R., & Alkhalaf, S. (2025). Secure edge-based IoMT framework for ICU monitoring with TinyML and post-quantum cryptography. Scientific Reports, 15(1), 36195.

Mishra, N., Habal, B. G. M., Garcia, P. S., & Garcia, M. B. (2024, June). Harnessing an AI-driven analytics model to optimize training and treatment in physical education for sports injury prevention. In *Proceedings of the 2024 8th international conference on education and multimedia technology* (pp. 309-315).

Zhou, D., Keogh, J. W., Ma, Y., Tong, R. K., Khan, A. R., & Jennings, N. R. (2025). Artificial intelligence in sport: A narrative review of applications, challenges and future trends. *Journal of Sports Sciences*, 1-16.

WAN, Y. (2026). Intelligent Algorithms in Enhancing Sports Performance: Theoretical Reconstruction, Technological Breakthroughs, and Future Challenges.

Khan, U. H., Qamar, A., Khan, R., Alawad, M. A., Alturise, F., & Hayat, M. (2025). TrustEdge: A Quantum-Safe, Self-Healing Framework for Federated TinyML in Critical ICU Monitoring. Internet of Things, 101855.

Chu, L. (2026). Big data-driven sports injury risk prediction model and preventive training strategy management framework. *Information Resources Management Journal (IRMJ)*, *39*(1), 1-19.

Luca, E. W. D., Dall'Ora, N., Giuliano, R., Della Valle, C., Di Cagno, A., Ferramosca, A., ... & Arcidiacono, G. (2026). A conceptual framework for athlete health using AIoT, wearables, and personalized performance intelligence. *Applied Sciences*, *16*(9), 4542.

Khan, U. H., Zia, A., Ali, H., Rahim, A., Tanveer, U., & Sher, K. F. (2025). Duplicate pull requests: Automated detection using s-bert and machine learning. *Spectrum of Engineering Sciences*, 991-1010.

Mateus, N., Abade, E., Coutinho, D., Gómez, M. Á., Peñas, C. L., & Sampaio, J. (2024). Empowering the sports scientist with artificial intelligence in training, performance, and health management. *Sensors*, *25*(1), 139.

Khan, U. H., Khan, R., Chelloug, S. A., Alturise, F., Khan, S., & Alkhalaf, S. (2026). HybridTrust: on-device federated learning with crypto-agile security for legacy and quantum-safe medical devices. Scientific Reports.

Musat, C. L., Mereuta, C., Nechita, A., Tutunaru, D., Voipan, A. E., Voipan, D., ... & Nechita, L. C. (2024). Diagnostic applications of AI in sports: a comprehensive review of injury risk prediction methods. *Diagnostics*, *14*(22), 2516.

Dong, G., Tan, X., Yuan, P., & Li, Y. (2026). Artificial intelligence and wearable sensors in sports injury risk prediction: current status and future perspectives: AI-wearable sports injury prediction. *Annals of Medicine*, *58*(1), 2658879.

Madrigal-Cerezo, R., Domínguez-Sanz, N., & Martín-Rodríguez, A. (2026). Wearable Biosensing and Machine Learning for Data-Driven Training and Coaching Support. *Biosensors*, *16*(2), 97.

Shah, M., Shah, A., Patel, K., Kshirsagar, A., Sanghvi, S., & Sojitra, V. (2025). Predictive analytics, strategic game analysis, and injury prevention in sports: the role of big data and artificial intelligence. *Machine Learning for Computational Science and Engineering*, *1*(1), 17.

Ahmad, B., Ali, S., Muneer, M., Fatima, A., & Abbas, N. (2025). Wind energy integration into the SAARC region: A comprehensive review of optimal wind sites for a sustainable super smart grid. Wind Energy, 3(2).

Guelmami, N., Fekih-Romdhane, F., Mechraoui, O., & Bragazzi, N. L. (2023). Injury prevention, optimized training and rehabilitation: how is AI reshaping the field of sports medicine. *New Asian Journal of Medicine*, *1*(1), 30-34.

Ahmad, B., Majeed, M. K., Soomro, A. A., & Jillani, S. A. (2026, January). Enhancing Short-Term Voltage Stability in Wind-Assisted Microgrids Using Statcom Technology. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Reis, F. J., Alaiti, R. K., Vallio, C. S., & Hespanhol, L. (2024). Artificial intelligence and Machine Learning approaches in sports: Concepts, applications, challenges, and future perspectives. *Brazilian journal of physical therapy*, *28*(3), 101083.

MD, D. A. J., & KOLLI, D. E. T. (2025). The future of sports training: Integrating artificial intelligence and wearable technology in performance enhancement. *TPM–Testing, Psychometrics, Methodology in Applied Psychology*, *32*(S2 (2025): Posted 09 June), 2145-2153.

Khan, U. H., Rahman, F., Mohammad, S. N., Rahat, A., & Abbas, S. S. (2026). FEDERATED TINYML WITH POST-QUANTUM SECURITY: QUANTIFYING THE PRIVACY–ROBUSTNESS TRILEMMA IN CRITICAL CARE IOMT. *Spectrum of Engineering Sciences*, *4*(3), 3945-3966.

Fyntikakis, E., Plakias, S., Tsatalas, T., Mina, M. A., Xenofondos, A., & Kokkotis, C. (2026). Convergence of Artificial Intelligence and Wearables in Strength Training and Performance Monitoring: A Scoping Review. *Applied Sciences*, *16*(7), 3565.

Wang, F., Jiang, S., & Li, J. (2025). The AI-Driven transformation in new materials manufacturing and the development of intelligent sports. *Applied Sciences*, *15*(10), 5667.

Chidambaram, S., Maheswaran, Y., Patel, K., Sounderajah, V., Hashimoto, D. A., Seastedt, K. P., ... & Darzi, A. (2022). Using artificial intelligence-enhanced sensing and wearable technology in sports medicine and performance optimisation. *Sensors*, *22*(18), 6920.

Hoseinzadeh, Z., Eskandarnejad, M., & Ghaderzadeh, M. (2026). Analysis of research structure and scientific trends in artificial intelligence, machine learning, and deep learning in football: a bibliometric approach. *International Journal of Intelligent Computing and Cybernetics*, 1-26.

Khan, U. H., Ali, H., Zia, A., Khan, H., Tanveer, U., & Bibi, S. (2025). PRIORITIZING PULL REQUESTS THROUGH DUAL PREDICTION OF ACCEPTANCE AND INTEGRATOR RESPONSE. *Spectrum of Engineering Sciences*, 1701-1715.

Menges, T. (2024). Advancing Endurance Sports with Artificial Intelligence: Application-Focused Perspectives. In *Artificial Intelligence in Sports, Movement, and Health* (pp. 31-48). Cham: Springer Nature Switzerland.

Jubair, H., Chowdhury, R. T., Akter, F., Sami, I. S., Nawal, N., Arshi, S. T., & Shil, P. (2025). Machine learning for real-time exercise correction and injury prevention: a systematic review. *Machine Learning*, *13*(1).

Nagorna, V., Mytko, A., Borysova, O., Potop, V., Petrenko, H., Zhyhailova, L., ... & Lorenzetti, S. (2024). Innovative technologies in sports games: A comprehensive investigation of theory and practice. *Journal of Physical Education and Sport*, *24*(3), 585-596.

Ahmad, B., Jillani, S. A., Rafique, M., & Soomro, A. A. (2026, January). Design and Monitoring of a Tree-Inspired Vertical Axis Wind Turbine for Efficient Urban Wind Utilization. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.

Zhao, H., Cheng, Y., & Gong, T. (2026). Artificial intelligence injury prediction model based on convolutional neural network: A new method to improve the efficiency of exercise rehabilitation. *Journal of Mechanics in Medicine and Biology*, *26*(04), 2640049.

Stetter, B. J., & Stein, T. (2024). Machine learning in biomechanics: Enhancing human movement analysis. In *Artificial Intelligence in Sports, Movement, and Health* (pp. 139-160). Cham: Springer Nature Switzerland.