Digital twin framework for injury risk prediction and management in competitive sports

Xiaojun Wang

Digital twin framework for injury risk prediction and management in competitive sports

Xiaojun Wang

Weinan Normal University,
Weinan, Shaanxi, 714000, China
Email: wxj804230@126.com

Abstract: This paper presents a digital twin (DT) framework aimed at predicting and managing injury risks in competitive sports. The proposed framework integrates biomechanical data, machine learning, and real-time monitoring to support athlete health and performance. DT technology enables detailed performance analysis and illness prediction through virtual athlete models. While previous studies have used wearable sensors and machine learning, they often lack cross-domain integration. This research addresses that gap by introducing a nine-stage DT development pipeline incorporating biomechanical data, PCA-based preprocessing, and musculoskeletal modelling. Validation through cross-validation and evolutionary data scenarios demonstrated the model's robustness. XG Boost achieved the highest injury prediction accuracy (78%), with key predictors including hamstring force and muscle stiffness. Biomechanical simulations revealed stress patterns consistent with physiological behaviour, supporting clinical relevance. Comprehensive decision-support systems remain scarce. This work contributes toward safer, more personalised sports environments using a holistic, data-driven DT approach.

Keywords: digital twin; DT; injury prediction; biomechanics; sports analytics; machine learning; musculoskeletal modelling; athlete monitoring; explainable AI; XAI; data-driven sports management; injury risk assessment.

Biographical notes: Xiaojun Wang graduated from Xi'an Sport University and received his Master's in Physical Education. He is currently teaching at Weinan Normal University as an Associate Professor, with a research focus on physical education teaching and training. He hosted and completed two scientific research projects of Shaanxi Provincial Department of Science and Technology, three projects of Shaanxi Provincial Sports Bureau, and published more than 15 papers.

1   Introduction

Digital twins (DTs) have started to blur the lines between the physical and virtual worlds. An electronic replica of a real-world item or procedure is known as a DT in the field of computer science (Laubenbacher and Mehrad, 2024). Although DTs have been around

for a long time, they are currently used in the real world and have several established uses. The medical field is just one of several modern-day applications of DTs; others include aerospace engineering and sports (Ferrari and Willcox, 2024). The ability to monitor surroundings in real-time and conduct numerous analyses is made possible by DTs, which are virtual representations of physical assets. Effective knowledge extraction from data requires genuine, real-time data, which data-making decisions provide. By letting users make digital versions of products before producing a physical one, it helps users optimise resources and prices in a number of sectors (Batty, 2024). The use of DTs in Industry 4.0 and, subsequently, Industry 5.0, can enhance manufacturing processes. Additionally, they can help optimise processes in the era of climate change by reducing carbon footprints (Hliš and Fister, 2024).

1.1   Defining the 'digital athlete'

A digital athlete uses data and technology to improve their performance and avoid injury. Their physical talents are enhanced and tracked through the utilisation of VR training, analytics driven by AI, and wearable gadgets (Alderson and Johnson, 2016). This article discusses the 'digital athlete' and how ideas like computer vision (CV), artificial intelligence (AI), and related ones help athletes perform at their best (Gabrišová et al., 2025). The advantages of becoming a digital athlete are many, and they include:

- By meticulously tracking and analysing data, digital athletes may pinpoint their weak spots and work to improve them, ultimately leading to higher performance. By utilising this vital knowledge, individuals can fine-tune their training routine, leading to notable advancements (Tokas and Kumar, 2024).

- Injury prevention for digital athletes has been transformed by technology. The ability to track one's physiological reaction to exercise allows for the early detection of any injury hazards. Athletes can prevent injuries before they happen with this cutting-edge gadget (Barricelli et al., 2020).

- The use of technology allows digital athletes to track their recuperation status and adjust their training program, allowing for a safe and quick return to competition.

With a focus on the consumer, we aim to implement DT technology as a means of managing mega-sport events (Jiang et al., 2021). Explored in this article are several aspects of how DT technology can change the character of modern athletic events. By improving fan experiences and engagement, DTs of sports venues help us understand how and why these events become more customer-centric. The next step is to promote DT technology as a fresh approach to managing mega-sport events at all points of contact with the consumer. Through a customer-centric examination of DTs and sports events, this essay will shed light on the complexities of using this new technology to improve the experiences of stakeholders and fans at significant sporting events. Finally, this article focuses on the client by outlining the advantages, disadvantages, opportunities, and challenges of DT technology as it pertains to athletic events (Bettega and Düggelin, 2023).

    The primary question that guides this research is: 'How (and why) do emerging technologies affect the customer experiences of sports spectators in the future through DTs of venues?'

    In addition, there are a couple of supplementary inquiries:

1 How can we trace the field's intellectual and historical evolution?

2 In terms of DT technology, what are the most significant developments and problems?

3 Who gains from the widespread use of DT technologies, and why are they essential for managing sporting events specifically?

4 With these trends and concerns, what are the advantages and disadvantages?

5 When it comes to sporting events, why are DT technologies being used so frequently, considering the environmental concerns around technology, spectators, and customer expectations?

In order to achieve this, we will analyse and describe the following points:

1 the history, development, and function of technological systems at sporting events

2 the factors that impact the deployment of technologies in sports entertainment, both now and in the future, from the spectators' point of view (Miehe and Waltersmann, 2021)

3 the factors that help improve the customer-centric approach to managing sports mega-events (Glebova et al., 2023)

4 an analysis of the DT's influence on event planning and execution

5 a look at how technological advancements are altering sporting events

6 detailing the function of DTs as event spectators in the future

7 a list of the top research and development priorities.

This article presents a theoretical framework by defining DTs as items in stadiums and by naming associated buildings and events.

The structure of this article is as follows: Section 2 provides a comprehensive literature review on the DT Framework for injury risk prediction. Section 3 outlines the proposed methodology for athlete management in competitive sports. Section 4 presents the results along with a discussion of their implications. Finally, Section 5 offers the conclusions drawn from the study.

2 Literature review

The advent of wearable sensors, video tracking, and machine learning has radically changed sports analytics. Thanks to these advancements, we can now look at physiological and biomechanical data as it happens. This shift from manual notational approaches to automated systems is a huge step forward. Many existing systems still prioritise isolated goals without integrating across domains; these include injury prediction, load monitoring, and tactical optimisation, among many others. In settings involving elite athletes, this dispersion reduces the potential for meaningful influence (Rico-González and Los Arcos, 2020). Furthermore, widespread problems such as poor data quality, unintelligible models, and unethical data governance still prevent widespread use. Make use of explainability methods like Shapley additive explanations

(SHAP) and local interpretable model-agnostic explanations (LIME), usually increases user confidence in the model's predictions. This is because these methods make model predictions seem more transparent and interpretable. Physiological data or GPS coordinates have also been employed in several studies as the basis for predictive models. However, decision-support systems that are thorough, explicable, and applicable across domains are still uncommon. To address these challenges and offer stakeholders real-time insights along with model transparency, Physiological, biomechanical, and tactical data are integrated into a unified big data analytics platform in three sport-specific use cases, as presented in this work (Mendes-Neves and Meireles, 2023).

2.1   Foundations

Data engineering, analytical modelling, and decision deployment are the three cornerstones of modern sports analytics. Working in tandem, they convert raw information into insights that can be used in peak performance settings (Mănescu, 2025). Data engineering can be defined as the practice of gathering, cleaning, and organising large datasets from many sources (such as user-generated content, live video streams, and wearable sensors). Structured data, like heart rates, semi-structured data, like videos, and unstructured data, like medical notes, all require systems to be able to sync, detect anomalies, and standardise schemas (Mundt, 2023). Analytical modelling, by processing processed data, obtains insights that are both predictive and prescriptive. The results of injury risk prediction, fatigue estimation, and tactical optimisation methods can be improved by integrating dimensionality reduction, supervised or reinforcement learning, and descriptive statistics. Dashboards and alarm systems that display information in real-time are the result of decision deployment. Visual dashboards (such as Grafana or Tableau) and mobile alerts are used to convert the model's output into actionable, context-aware insights tailored for coaches, analysts, and medical staff.

2.1.1   Sports-related concussion assessment

There have been more in-depth conversations about how to effectively detect and manage SRC cases due to an increase in their appearance at A&E and ED, with some even calling for the creation of national guidelines. Draw attention to the fact that the primary goal of SRC evaluation in the emergency room is to sort the patient or athlete according to the severity of brain injury, with the goal of determining if immediate action, such as surgery, is necessary (Celik et al., 2020). While that method could potentially increase productivity, it fails to evaluate numerous other modest SRC impairments, including visual, cognitive, and motor/functional issues (such as balance and gait). Therefore, many of the present SRC evaluations are simplistic, one-sided snapshots that fail to take into account the complexity and diversity of the human condition. Providing players with information about warning signs and symptoms and/or informational booklets about head injuries constitutes the majority of post-discharge management. As an added downside, until returning to full contact training, there is frequently no physician evaluation or follow-up outside of professional settings. So far, self-reported symptoms and measurements have been the backbone of SRC treatment and rehabilitation regimens for determining play readiness. Thus, SRC prognosis and recovery durations fluctuate considerably between genders and age groups.

2.1.2 Cognitive: routine clinical approaches

When evaluating cognitive function in contexts other than athletics, it is common practice to conduct in-depth interviews with patients to learn about their medical, educational, and social backgrounds. Sport-related concussion (SRC) primarily affects key cognitive functions such as working memory, executive function, and short-term memory (Stuart et al., 2020). Among the tests that were once given on paper were the standardised assessment of concussion (SAC), which is now a component of SCAT5, the digit span (both forward and backward), and the short-blessed test. One major limitation of these tests, despite their extensive clinical use, is that they cannot be automatically or instantly compared across several individuals or periods due to the requirement for manual score calculation. Thankfully, these constraints have been somewhat circumvented by the development of computerised neurocognitive testing (Powell et al., 2021).

2.1.3 Cognitive: digital approaches, computerised programs

Automated data collecting and processing, randomisation of test sessions, and the use of objective cognitive metrics (such as response time calculation) are some of the benefits of digital-based cognitive exams compared to pen-and-paper procedures. The ImPACT was developed to evaluate verbal and visual memory, visual-motor speed, reaction time, and the immediate post-concussion assessment and cognitive test. You can see an example of a scalable computerised neuropsychological instrument here. The ImPACT test includes a demographic assessment of both staff and players as well as an evaluation of symptoms experienced after a concussion (Jones et al., 2020). According to research, the impact is sensitive during the acute phase (the first few days after a concussion). It can identify changes in visual and verbal memory, as well as prolonged reaction times. The evidence for assessing cognition after a subacute or chronic concussion is inconsistent. In accordance with the global consensus on concussion, 'Tests should not be seen as the sole basis for the management of decisions'.

3 Materials and methodology

Figure 1 outlines a nine-stage pipeline for constructing and deploying a DT system in biomechanical and physiological domains. It starts with dataset collection and progresses through preprocessing, physical modelling, and machine learning model training. A DT architecture is then designed and integrated with real-time monitoring and validation. Feedback loops enable continuous model refinement and performance optimisation.

3.1 Dataset for biomechanical tests

The final dataset included assessments from 110 male professional soccer players from three of Chile's most famous first-division clubs. Age, weight, height, anthropometric measurements, results of various biomechanical tests, and team positions were among the 19 characteristics taken into account for each participant (Calderón-Díaz and Silvestre Aguirre, 2024). All evaluations and testing were conducted by kinesiologists from the Innovation Center's Biomechanics Laboratory in Santiago, Chile. The MEDS Clinic includes the laboratory. Injuries sustained within the past three months or a BMI

below 24 disqualified you from participation. We checked that everyone who needed to be included or excluded had provided their informed consent before we started collecting data. Ensuring the confidentiality of the data, the MEDS clinic anonymously submitted the data matrix.

Figure 1   System optimisation and real-time monitoring made possible by a DT development pipeline that integrates data, modelling, and machine learning (see online version for colours)
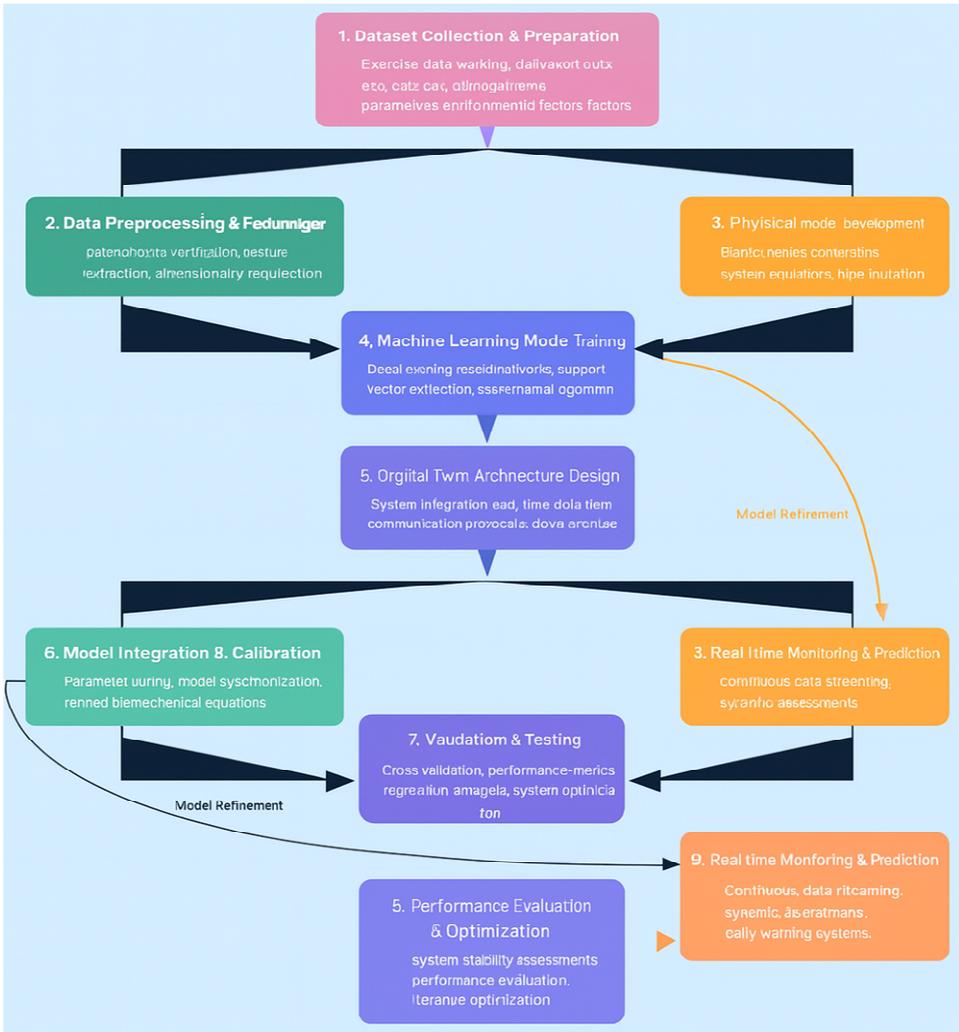

3.2   *Biomechanical tests*

In the eccentric asymmetry force test, often known as the Nordic hamstring test, test subjects kneel while supporting their trunks with their hands [Figure 2(a)]. To keep the heels from coming off the ground while you work out, assistant or load cells must be used. The eccentric activation of the hamstring muscles can be measured using load cells.

The most significant eccentric force (in Newtons) applied to the right and left hamstrings, respectively, are the two metrics that emerge from this examination. Figure 2(a) shows aquadriceps chair that participants sit on with load cells attached to measure eccentric quadriceps activity (Adolfo et al., 2021). The maximum force quadriceps test is another name for this exercise. The test measures your left quadriceps force (N) and eccentric force (N), two critical metrics. Clinically, a one-leg bridge test can reveal how vulnerable a patient is to hamstring tears. Assuming a supine position on the floor, the participant has to place the assigned leg's heel inside a 60 cm high box. The participant is required to lift their glutes off the floor by pushing with their heel while crossing their hands over their chest. See Figure 2(b) for a visual representation of how each repetition involves touching the ground and then immediately lifting the glutes again without resting. The results of this test will show you how many times you should do the right and left legs separately.

Myotonometry is a digital palpation technique that measures the stiffness of superficial skeletal muscles in an objective and non-invasive way. The Myoton PRO gadget is used to measure the hamstring muscles, as shown in Figure 2(c). On both ends, you can see the S-stiffness (N/m), which is a metric for how resistant something is to stretching or compression that would otherwise deform its structure or tissue. Using the Bosco method to measure vertical leap: An examination of the morphophysiology, functional qualities, and neuromuscular characteristics of the lower limb extensor muscles will be conducted through a battery of vertical jumps. These jumps will measure mechanical jump power, height attained, elastic energy utilisation, fatigue resistance, and myotatic reflex. Using three force platform leaps, the Bosco test is conducted. The squat jump, Abalakov jump, and countermovement jump are illustrated in Figure 2(d), together with the corresponding data and their execution. The squat jump and the Abalakov jump are both bipodal and unipodal, respectively.

Figure 2 Biomechanical evaluation protocol (see online version for colours)
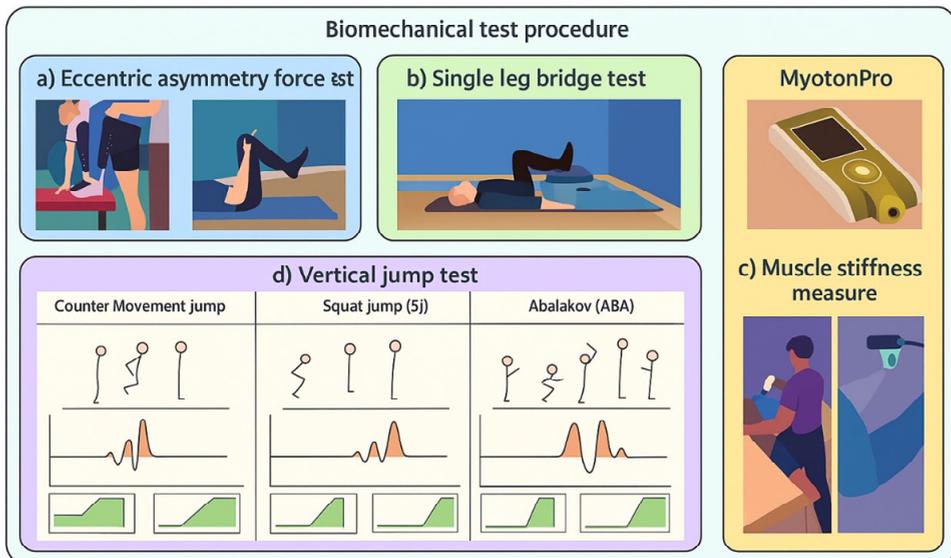
3.3   Data-cleaning applications

Nowadays, there is an exponential growth in the amount of data collected, but the reliability of that data is deteriorating due to the fact that much of it is contaminated. There are a number of data-cleaning approaches that can fix this problem, but dealing with big data makes data cleansing much more of a challenge. Data cleansing (or data cleaning in short) is an old field of study. By identifying and removing mistakes and discrepancies, it hopes to improve data quality (Rocha et al., 2025). Extensive data cleansing and classical data cleansing are currently two separate subfields within data cleansing. Conventional data cleansing methods, such as Intelliclean and Potter's Wheel, struggle with massive datasets.

3.3.1   Principal component analysis-sample reduction process

Minimising the number of variables while preserving the maximum degree of variation is the goal of principal component analysis (PCA) in a dataset containing several related variables. PCA keeps most of the original variables' variance while changing them into a new set of independent variables. One way to rank new orthogonal, linearly independent dimensions or measurements is by looking at their data variance. Since the significance of the most crucial central axis indicates the degree of dispersion or variance in the data, it is displayed first. To understand this, you must have a strong background in PCA, variance, covariance, eigenvalues, and eigenvectors. Exhibiting a strong positive correlation, as seen in Figure 3, we can deduce that X and Y are highly connected; that is, as X grows, Y likewise increases. A negative covariance shows the inverse relationship. X and Y are unrelated if their covariance is zero.

**Figure 3**   Covariance and its values (see online version for colours)
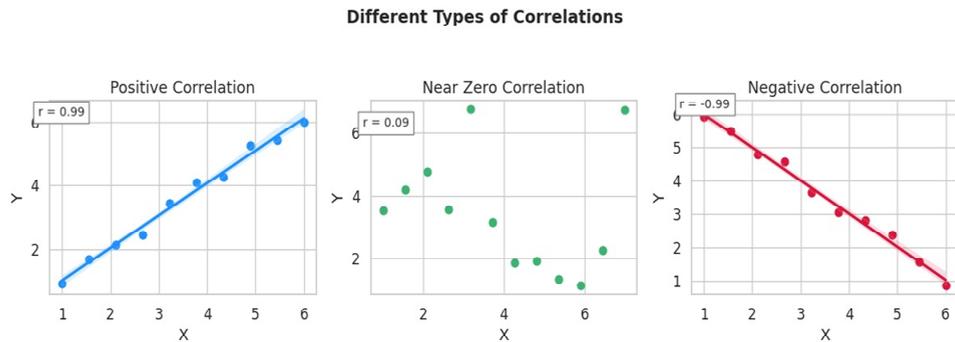
One great way to get to know the patterns in multidimensional datasets is through data visualisation. When data is plotted on a two-dimensional plane with the x- and y-axes, it is easy to see the pattern. Still, when dealing with multidimensional data, which has many features to take into account and analysis to compute, it becomes much more difficult to visualise. To prepare for PCA, you must first examine the features' contributions to the dataset as a whole using eigenvector analysis, which involves finding patterns in the dataset in order to distribute the data over all dimensions. Afterwards, it retains the feature columns with the highest eigenvalues and performs dimension reduction. Figure 4 is utilised to illustrate the effectiveness of PCA. This figure projects a collection of data

from multiple dimensions into a two-dimensional environment. Finding a linear correlation between data points might be difficult because of the high dimensionality of the data. After the points are transformed into column vectors, they are combined into a matrix M. After that, M is subjected to PCA in accordance with equation (1):

$$PC(M) = W \times V \tag{1}$$

where the matching eigenvalues are denoted by W and the eigenvectors by V.

Figure 4   Data points are projected at two angles (see online version for colours)
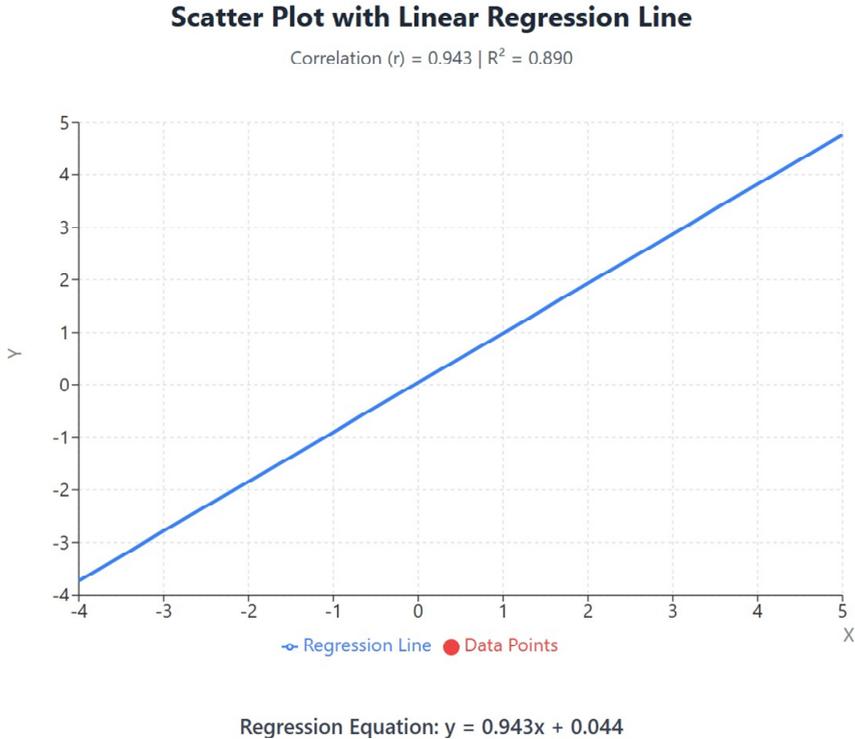
The process of generating $XPC1$ is illustrated in Figure 5 by comparing the top k eigenvalues in W with the covariant eigenvalues V. Then, these values are chosen to reflect the data points best. The red lines that run through the spots that show the correlation amongst them represent the primary axes ($PC1$ and $PC2$).

Lastly, the most crucial aspect of the PCA-sample reduction method is:

$$PCA_{SRP}(M) = \left| M \times V_{PC_1} \right| \tag{2}$$

In dataset M, the loading score is represented as $PCMTR(M)$, and in matrix M, the $PC1$ eigenvectors of the samples are $VPC1$.

$$S = \sum_{0}^{D} \begin{cases} 1 & \text{if } PCA_{SRP}(M) > Sc \times PCA_{SRP}(M)_{highest} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

In a dataset with D samples (including random and noise samples), assuming that $S \le D$, S is the quantity of samples that are higher than the set-biased Sc loading score.

$$passrate = S / D \tag{4}$$

In dataset M, the *kayyfaZb* represents the rate of samples cleaned using PCA-SRP, and its calculation is done using equation (3). The new process dataset DPS will accept samples over the stated threshold and reject samples below it, according to equation (4).

Figure 5   Data point projections into the two freshly built dimensions (see online version for colours)
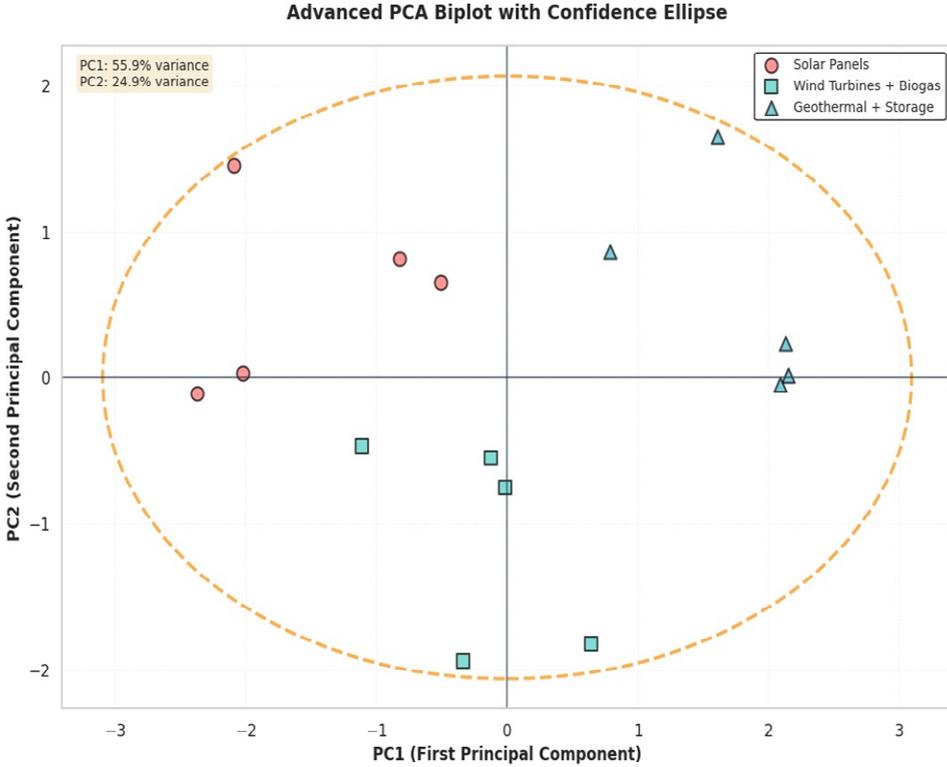
### 3.4   *Musculoskeletal model*

Aalborg, Denmark-based Any Body Technology's Any Body Modelling System (AMS), version 8.0, was used to conduct the musculoskeletal system investigations. Our primary resource for this study was the BVH X sens model, which is a comprehensive representation suitable for gait analysis and can be found in the Any Body Managed Model Repository (AMMR, version 3.0.4). This AMMR model, grounded in anatomical data derived from cadaver research, enables the execution of biomechanical simulations. Using data from motion capture to guide joint kinematics, parts of the lower and upper extremities, and key muscle groups, are all part of this model (Rossi and Pappalardo, 2022).

A number of adjustments were made according to the topic at hand:

- *Lab specific data*: Verifying proper detection of ground contact (shown by red spheres) is the responsibility of this component. We adjusted the ground contact threshold to 0.045 for both legs because it significantly affects the forces pushing on the knee joint.

- *Subject specific data*: Individual characteristics inform the calibration of the subject's height and weight parameters.

- *Trial specific data*: A frame count of 10 and a frame count of 300 make up the examined trial. When the subjects' gait cycles were positive, we only used the frames that corresponded to forward walking for this investigation. This meant that the musculoskeletal modelling could smooth out the bumps in the road brought on by abrupt changes in gait pattern, such as turning to face a different direction while walking.

- *Any Mocap model*: provided the necessary power to the skeleton model through the analysis of the motion capture data.

These adjustments allowed for the model to be fine-tuned for gait analysis, taking into account the participant's motion and loading conditions. Following the building of the customised MS model utilising the collected kinematic data, inverse dynamics was used to anticipate the internal joint forces and moments. Inverse dynamics considers the body's mass distribution and inertial characteristics to compute the net joint forces and moments when a measurable motion is being calculated. In order to determine the joint response forces and moments during the gait cycle, the Any Body program applied the Newton-Euler equations of motion to the joint angles obtained from motion capture (Figure 6).

Figure 6 Block diagram illustrating inverse dynamics for estimating joint forces and moments from motion and inertial data (see online version for colours)
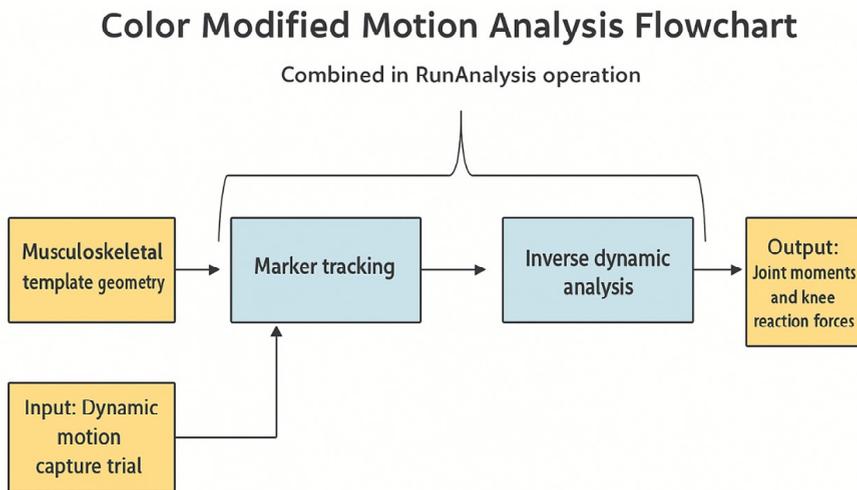
With ligament integration scheduled for a subsequent step, the initial phase solely focused on the femur and tibia bone structures. To facilitate integration between the Any Body MS model and the FE knee model, the femur and tibia geometries were isolated

and exported as stereolithography (STL) files. The design of each structure was included in a new file that was produced within the Lab Specific Data section. In order to transform millimeters into meters, the Any Draw Surf class was used to draw and show the geometries, using a scaling factor of 0.001. The STL models were adjusted to match the human anatomy in Any Body by applying a rigid-body transformation with the AnyFunTransform3DLin2 object. We used VTK_LANDMARK_RIGIDBODY as our transformation mode as it only performed rotation and translation; in this way, we avoided applying any scaling at all. Because of these anatomical landmarks, the musculoskeletal model was able to be spatially registered for the first time with the precise knee geometries:

Point 0   STL-file reference points, including the medial and lateral femoral epicondyles, midway between them, the femur's knee joint anatomy, and the tibia's tenderness.

Point 1   The anatomical hierarchy of the Any Body model defines the target landmarks.

The current musculoskeletal model and the customised knee geometries are typically accurately spatially aligned using this procedure. But when I reloaded the model, the bones seemed like they were partially fused in the model view (Figure 7). It appears that a more sophisticated scaling method could be necessary to enhance the alignment and morphological fitting. This could involve utilising transformation functions like AnyFunTransform3DRBF or AnyFunTransform3DSTL.

Figure 7   The femur and tibia's STL geometries are aligned and transformed using the main anatomical reference points, which include the epicondyles, the knee joint anatomy frame, and the tibial tuberosity (see online version for colours)
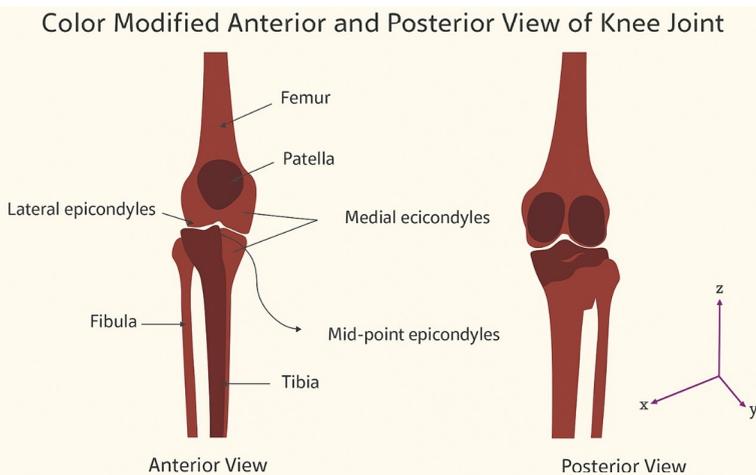
Note: After alignment, the Any Body Model View shows that these bones have fused somewhat, as seen in the image as well.

To specify how the forces acting on the tibia are extracted, a piece of code is inserted in the Lab Specific Data. Any file. At first, we measured the right tibia's forces and moments, which included those produced by the knee's joint reflexes and muscles like the psoas major. The programming also takes inertia and gravity into consideration. Translating the retrieved forces and moments into the anatomical reference frame of the

tibia ensures their proper orientation for usage in subsequent simulations, like those in Abaqus. The three-dimensional moments (M) and net knee forces (F) acting on the right tibia were derived using inverse dynamics analysis. These forces are documented in a CSV file for future research. Keep in mind that the simulations only account for the right knee when extracting forces and moments; this is because the 3D model utilised corresponds to that particular joint.

3.5   Train and test

3.5.1   Validation

It was common practice to divide the dataset into three separate datasets before evaluating the machine learning model. The first step is to fit the model's parameters using a training dataset, which is a collection of samples. As the model's hyper parameters are fine-tuned, the accuracy of the model's predictions is checked against a validation set. Last but not least, the test set is a dataset that is utilised to assess the final model impartially (Glebova et al., 2023). The ten-fold cross-validation method has been the de facto standard for validating injury forecasts in the aforementioned literature. Instead, they proposed an innovative validation approach that is based on the kind of challenges that soccer clubs encounter during the season. Regularly, the dataset is used to construct a training set and a validation set using the cross-validation technique. To ensure the accuracy of the model, an additional set of test data that does not use cross-validation must be used. A stratified split technique is essential for cross-validation to guarantee that classes are distributed uniformly throughout folds. It is expected that the distribution of courses within each fold aligns with the distribution throughout the entire dataset. Finally, we looked for previous studies that evaluated predictive capacity on a test set, but we could not find any. Week n was spent training and validating their models.

Tests were conducted in week $n + 1$. Plus, they used the season to train and validate their models, with the following year being used to evaluate the fitted models. Researchers need to conduct further studies to determine how accurate the damage predictions are, as almost all of the models suggested in the literature have not been tested. The classifiers' injury prediction skills can be tested using data that is not currently part of the dataset used to train and validate the ML models. This is made possible by the proposed evolutionary scenario. The key advantage of the evolutionary scenario, which allows for a more comprehensive examination of the model, is the capacity to forecast the forthcoming soccer season. By increasing the sample size with additional injury and no-injury cases, you can modify your model training as the season continues, contrary to the validation strategy given.

Allow me to explain: With data going all the way up to day n, a soccer club is attempting to calculate the odds of player injuries happening between $n + 1$ and $n + i$. To train and cross-validate machine learning models, one uses the training set, which consists of data up to the current day. In order to forecast the probability of harm, the models are then tested using additional data collected on days $n + 1$, $n + 2$, $n + i$. Researchers interested in assessing their model's predicted ability in a real-world setting with daily data generation must use this test approach.

3.5.2   Data processing for each training fold

Every training fold in the evolutive scenario technique and the cross-validation method incorporates data preprocessing. Since injury cases are significantly less common than non-injury ones, it is helpful to use a sampling approach to ensure that the target classes are balanced. The machine learning model is made more understandable and easier to work with by narrowing the feature space and choosing the most helpful traits for injury prediction, mentioned.

3.5.3   Sampling

Data analysis techniques that alter the class distribution of a dataset include over- and undersampling. More instances from the minority class are added through oversampling, and fewer cases from the majority class are decreased through undersampling, in order to attain statistical parity between the two groups. The injury dataset only includes around 2% of all observations, so an oversampling strategy is needed to make sure that both groups are represented fairly. Models trained using the oversampling method outperform those without it when it comes to damage prediction. The synthetic minority over-sampling approach (SMOTE) was extensively utilised in earlier studies. Within the feature space of dimensions, this method takes into account the k-nearest neighbours of a data point. An artificial data point is created by the SMOTE method using the vector of the difference between the current data point and one of those k neighbours. The next step in developing new simulated cases is to take the vector, add it to the present data point, and then multiply the result by a random number between zero and one. A few earlier studies found that using such an oversampling strategy improved the accuracy of damage predictions. The adaptive synthetic sampling approach (ADASYN) was employed instead, which is a variant of the SMOTE algorithm.

The SMOTE algorithm is the basis for this method, with the exception of a single key detail. For duping, ADASYN will choose a specific point that is not in a homogeneous neighbourhood. Furthermore, Carey and colleagues examine the impact of various sampling methods on injury prediction. They discover that while SMOTE does not impact performance, under-sampling the non-injury class lowers model performance. The authors of that study suggested that under-sampling might not be the best method for injury predictions because it could lead to information loss by excluding a large number of samples that did not involve injuries. To add insult to injury, SMOTE may work for patterns of injuries but will not do much good if fresh ailments do not resemble old ones. However, the same issue may arise if no oversampling strategy is used, since new injuries may vary from prior ones, making prediction impossible.

3.5.4   Feature selection

The term 'feature selection' refers to the steps used to narrow down the pool of potential features associated with injury prediction output to a manageable set. By flattening the feature space, this method speeds up the training of machine learning models, improves their interpretability, and decreases the likelihood of multicollinearity problems. As far as we are aware, this method was employed in a single study. Made use of the RFECV technique to identify the most critical features for injury prediction. By iteratively examining progressively smaller sets of features, RFE aims to pick features. It all starts with training the estimator on the complete training dataset to determine the relative

relevance of each feature. The next step is to remove the features that are not crucial from the present set. After that, the subset of features undergoes this process recursively until the ideal feature count is determined. By applying RFE to a separate portion of the training dataset, cross-validation helps to choose the most significant and consistent features from each fold. In order to ensure that the training set is not prone to overfitting while evaluating the goodness of the model, feature selection should be carried out.

3.5.5 Hyper parameters fit on validation set

Can the best machine learning hyper parameters be found using the validation set to improve injury prediction? Because they govern the model's general behaviour, hyper parameter adjustment is critical for machine learning. It is possible to alter the hyper parameters of any machine learning model. Grid Search is one of the earliest and most well-known methods for hyper parameter optimisation. Only two publications showed that injury prognosis models performed better after modifying the hyper parameters. This approach explores every conceivable combination of input values for the hyper parameters. By employing this approach, we may determine which values of the parameter search space's hyper parameters work best for each model. The processing resources needed for this are typically higher. The Random Search is another method for finding the optimal hyper parameters. Aside from testing each potential combination of hyper parameters, it is identical to the grid search method. The sole difference is that it uses a random technique. Since the second method uses less computing power, it can be pretty successful in practice; nevertheless, it does not necessarily produce the same ideal outcome as grid search.

3.6 DTs in sports event management: an explanation principles supporting the integration, diffusion, and upkeep of DT

A trustworthy framework and relevant theoretical constructs are required to explain the phenomena as they evolve. The data we have allows us to put forward a model that suggests how DTs could improve sports events in many ways, including:

1 audience participation

2 operational effectiveness

3 environmental friendliness, all of which could lead to more customer-focused experiences overall (Figure 8).

Consequently, we can suggest that sporting event organisers use DT technology to build virtual stadiums that are highly accurate representations of the actual thing, giving fans a more immersive and tailored experience. This, in turn, makes fans happier, more invested, and more devoted. Our findings also lend credence to the idea that DT technology, by combining real-time data with simulations, can optimise venue operations. Event organisers can cut expenses by using this technology to boost operational efficiency, simplify logistical operations, and better distribute resources. Also, it shows how DTs might help with environmental concerns during sporting events. As a result of DTs' ability to reduce carbon emissions and the necessity for physical attendance, among other sustainability-related considerations, sports consumption could become more ecologically conscientious. Therefore, the 'event digital reflection'

(Figure 8) highlights the potential impact of DT technology on the organisation, upkeep, and enjoyment of sporting events by spectators. Personalised, captivating, and environmentally friendly experiences at sporting events can be achieved by combining customer-centric strategies with cutting-edge data analytics and immersive technology.

Figure 8   The three cornerstones of DTs: digital reflection on sporting events (see online version for colours)
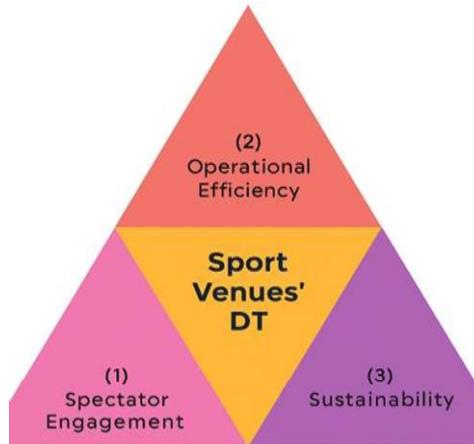
4   Results

In Table 1, you can see the results of 35 ML algorithms' testing, as well as the model's configuration and description. Figure 9 displays the accuracy values of all the models. The XG Boost method, with an accuracy of 78%, outperformed the others, with accuracies of 70% or higher. The next generation of algorithms included support vector machines, decision trees, logistic regression kernels, and kernel neural networks.

Maximal left hamstring strength, maximal right biceps femoris stiffness, and maximal semitendinosus stiffness were the most significant outcomes of XGBoost, as shown in Table 1. The most frequently repeated categorisation variable among the 30 iterations was given the most weight. Figure 4 shows that the maximal hamstring force is the most crucial attribute. Just like in Table 3, this case makes use of the same variables that were utilised to display the feature significance graph of the top-performing model.

Table 1   Key characteristics extracted by XGBoost (those that recurred most frequently throughout the 30 iterations)

Feature description	Count of occurrences
Peak force – left hamstring	28
Muscle stiffness – right biceps femoris	28
Muscle stiffness – right semitendinosus	24
Peak force – right quadriceps	21
Eccentric strength – hamstring muscles	17
Participant age	16

Figure 9   The testing accuracy of ML models is compared (using 10 k-folds) (see online version for colours)
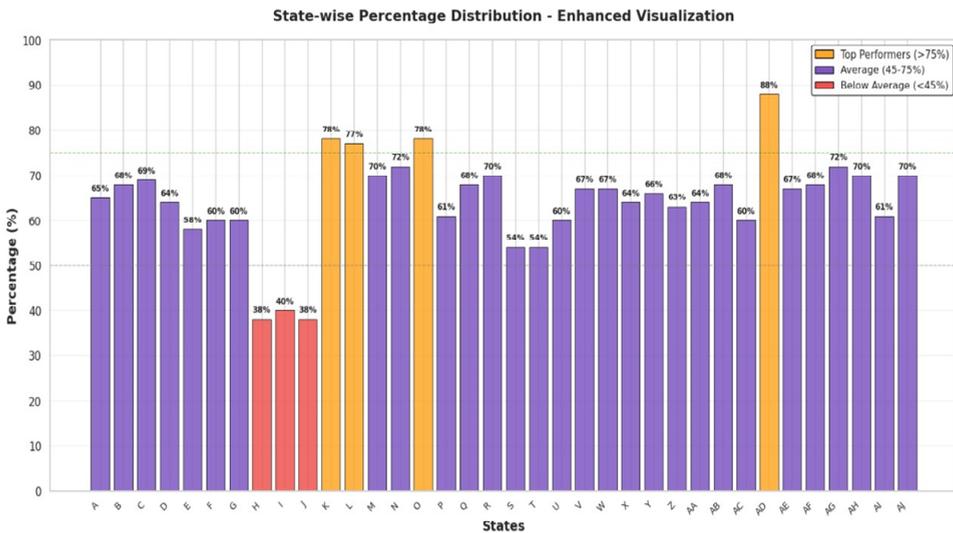

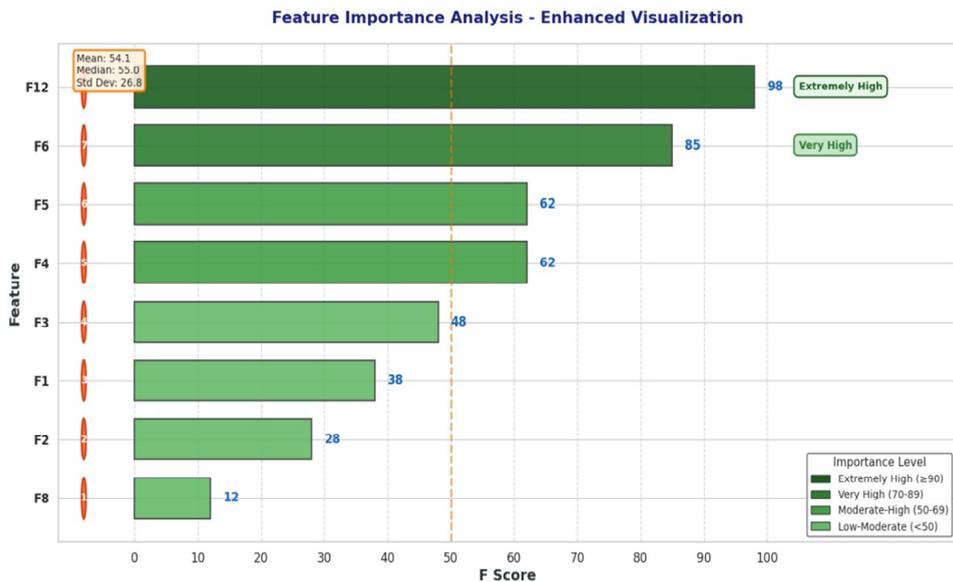**Figure 10**   Feature importance (see online version for colours)

Figure 10 demonstrates that the classifier with the lowest effectiveness is the one used with the cancer patients' dataset. To see how it would react to a 100% increase in random samples, we injected it with a spike. The result of PCA-SRP steadily lowers as the random samples increase, as shown in Figure 11.

Validation accuracy up to 100% was achieved by both the ANN + PCA(SRP) and ANN alone, with selectivities (Sc) of 98% and 88.2%, respectively, when exposed to

additional sounds (Figure 12). Additionally, by switching to 98% selectivity (*Sc*) from 88%, it has demonstrated a respectable improvement in accuracy.

Figure 11   PCA cleaned data percentage (%) (see online version for colours)
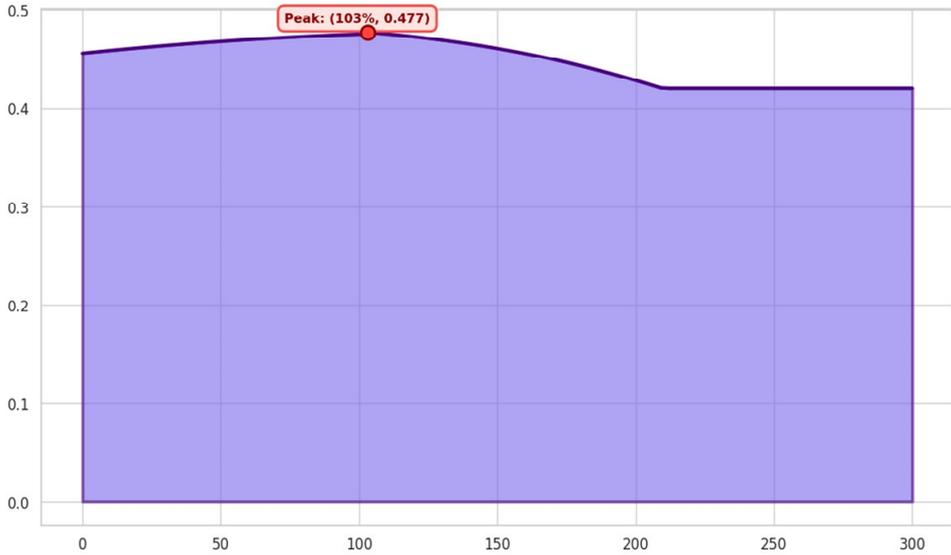

Figure 12   +Response to noise in a dataset of cancer patients (*Sc* = 88%) (see online version for colours)
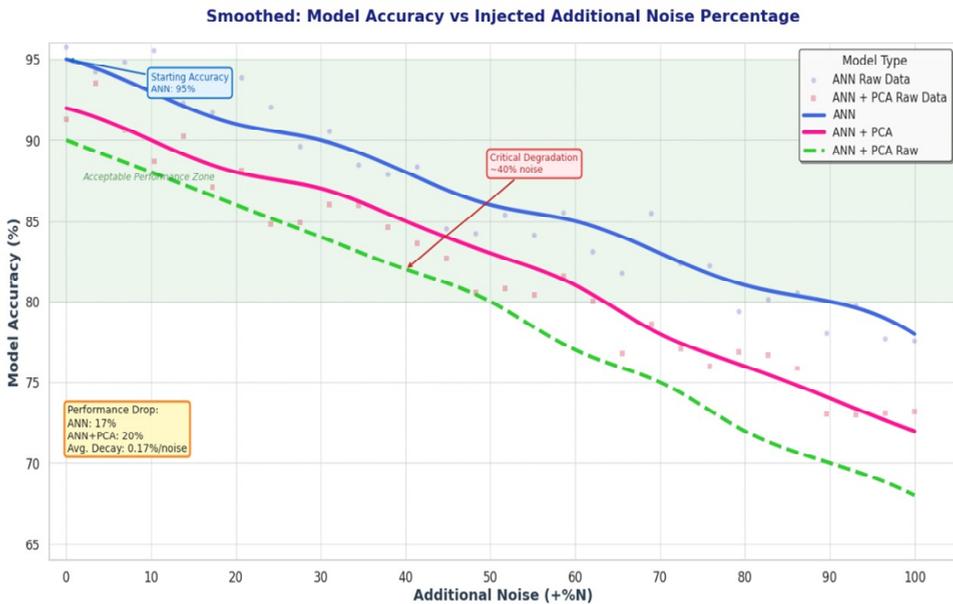
Figure 13 Views of the model's front (left) and back (right) sides reveal its maximum primary stress distribution (see online version for colours)
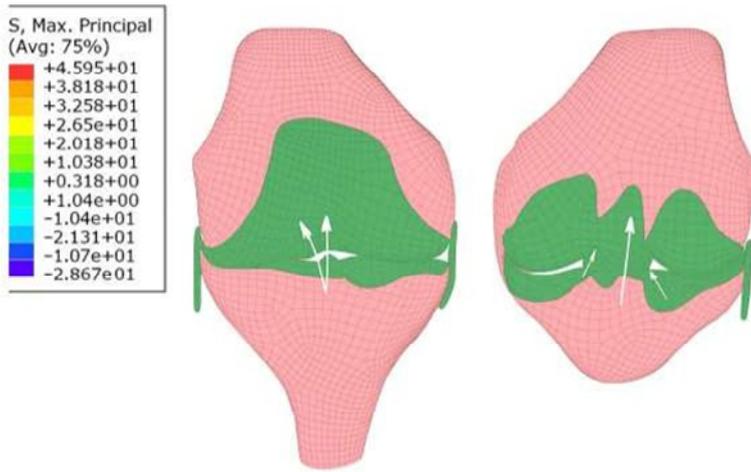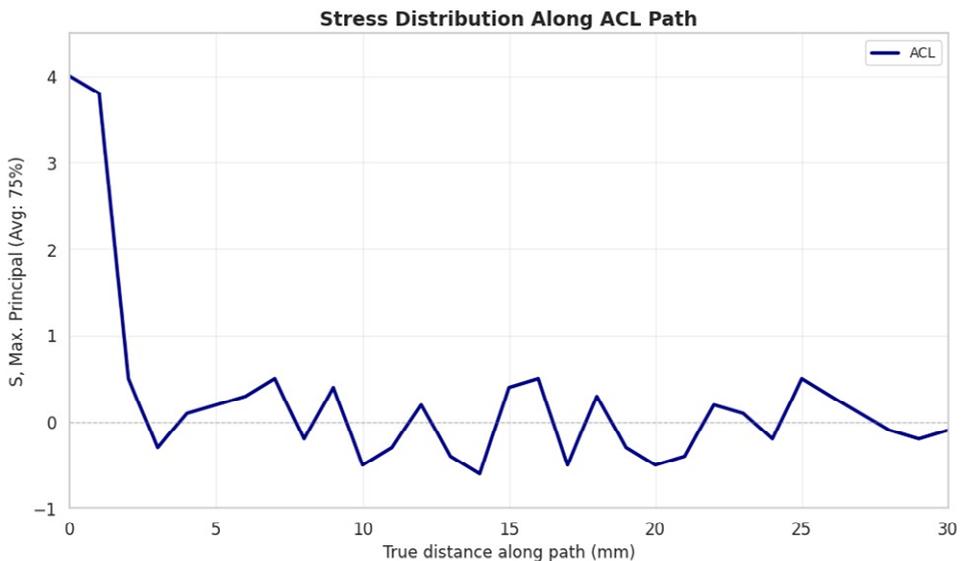

Figure 14 How the stress distribution along the ACL is affected by the actual distance from the femoral to the tibial insertion (see online version for colours)


The data that is highly valuable has been saved and will remain accurate up until a certain additional noise point. Even though some data was lost, the predictive power is still there up until that time. With high Sc values, the $ANN + PaMSRP$ cancer patient classifications continue to function at the highest level. By progressively minimising the fall in classification issue accuracy due to the rapid rise in system noise, the methodology provides a significant advantage. A large number of high Sc values are required to solve the ANN classification problem, whereas low ones are disregarded. According to the findings in Figures 12 and 13, as stated in, $ScANN$ is the superior value. With an IKDC

score of 98.9 and a KOOS score of 98, the participant's knee function was exceptional, demonstrating excellent levels of recovery with minimal functional damage. Following this, the Abaqus biomechanical simulation study will have a firm foundation in a clinical setting thanks to these findings. Figure 14 shows the distribution of maximum primary stresses from both the anterior and posterior views of the model, which summarises the results.

Dark blue represents stress values around –28.67 MPa, whereas red represents stress values around +44.95 MPa. Both views show that the medial and lateral tibial plateaus, where the green tonalities are most pronounced, are subject to relatively light compressive forces, consistent with the joint's ability to transmit axial loads. The considerably more verdant regions around the intercondylar region, which are probably linked to contact pressure peaks or ligament insertions, seem to be under a great deal of tensile stress. Results for stress in the tibial plateau area from the simulation were around 2 to 8 MPa, which is in good agreement with the 2 to 6 MPa mentioned for usual walking. This can mean that the patient's joint is fully functional again, able to distribute loads as it usually would when walking, thanks to the operation. To capture the truth of how joints work, though, a deeper examination of muscular forces is required. Suppose we want to make inferences about knee function after surgery that is applicable to a broader population. In that case, we need to increase the size of the sample and include individuals with different characteristics. After running the simulation, the stress fluctuation along the actual ACL path was presented on the graph in Figure 14.

Figure 15   Both the front and back images show the total magnitude of the knee joint's displacement (U, Magnitude) when subjected to loading conditions (see online version for colours)
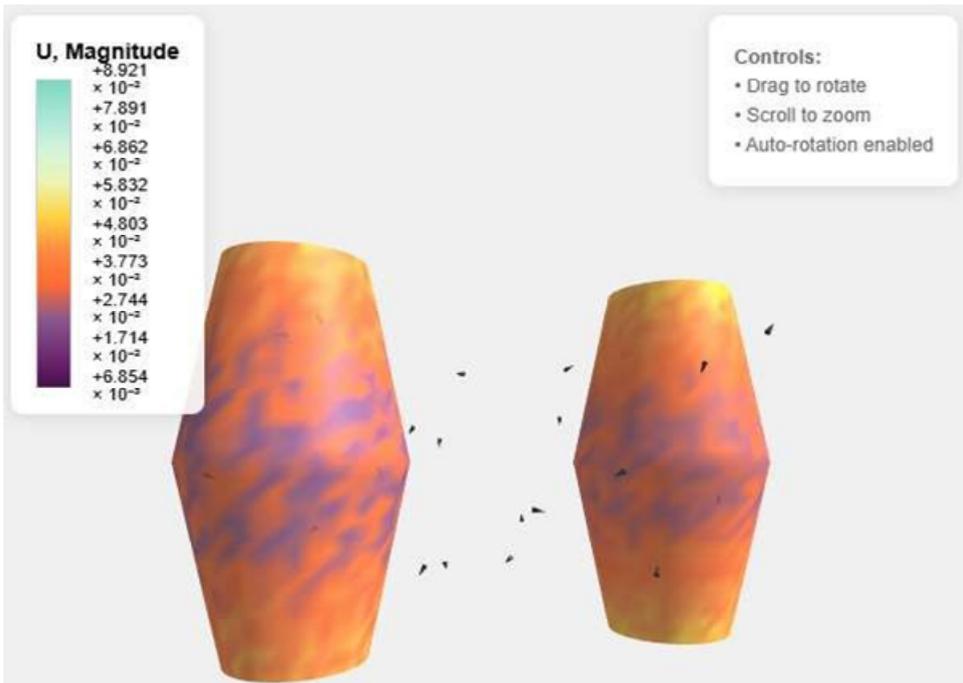
Indicating an area of substantial tensile loading close to the femoral insertion, there is a stress peak at the beginning of the path, close to 0 mm, which reaches approximately 1.45 MPa. These findings are consistent with those of, who also discovered that the anterior cruciate ligament (ACL) attaches to the femur and that landing is the most stressful time for the ligament. The stress falls precipitously to approximately 0.4 MPa immediately following this peak, and it continues to decline thereafter. Starting at approximately 5 mm and going onward, the stress levels are continuously low, varying between 0.0 and 0.2 MPa. Notable is the acceleration to a negative stress value of –3.0 MPa, which may refer to a transition zone at the fixation point or localised compression. When the femur is immobilised and the tibia is loaded, the area farther away from the tibia experiences less stress or even slight unloading. In contrast, the area closest to the femur, the proximal region of the ACL, appears to bear a significantly heavier load, according to the uneven distribution of stress. The ACL's curving anatomical course and the direction of applied force influence its usual mechanical performance. The results show this. Accompanying the stress analysis was a displacement investigation that assessed the knee joint's mechanical response to the imposed loading conditions. A comprehensive analysis of the joint structure's total displacement (U), including all directional components, is shown in Figure 15.

Figure 16 Measurement of displacement along a trajectory starting at the distal tibia and ending at the area where the ACL is inserted (see online version for colours)
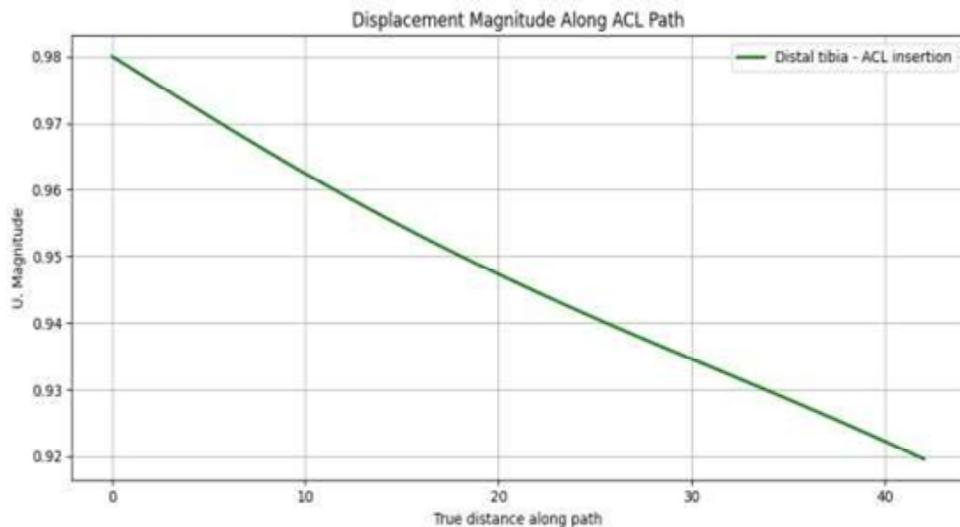
In entirely confined places, the colour scale reads 0 mm (dark blue), whereas in areas of highest movement, it reads about 1.068 mm (red). The displacement values are shown in millimetres (mm). Because of the boundary conditions, the femur stays relatively still, whereas the distal portion of the tibia shows the most significant displacement values. This is because this is the area that was loaded in the simulation. The lower and middle tibial regions, along with the soft tissues surrounding them, display intermediate displacement values, which are shown by colours ranging from green to yellow. The gradient between the unloaded femoral area and the loaded tibial extremity is the mechanical change, and deformation is distributed equally over the joint. As expected

from the knee's biomechanics, this displacement pattern confirms the mechanical effect of the loading scenario, which causes the tibia to move mostly lower and anteriorly in relation to the femur. See the displacement along the tibial structure in Figure 16 for a better understanding of the localised deformation behaviour under stress. Under the given boundary circumstances, this method sheds light on the tibial structure's displacement fluctuation.

As the path moves closer to its target, the displacement graph shows a steady and gradual decrease. At the far end of the joint, where the force is delivered, the maximal displacement is around 0.99 mm, but it drops to about 0.83 mm close to the ACL insertion. Consistent with these displacement values, FE models of the knee often show displacements ranging from half a millimetre to two millimetres, depending on the magnitude of the applied load and the tissue material properties. In particular, they noted a 1.10 mm distal tibial displacement when subjected to the identical stress conditions. The existing model is supported by the fact that there is only a 0.11 mm difference between it and the current simulation, indicating an excellent agreement with earlier FE analyses and a tiny variation. A confusion matrix, often known as an error matrix, is a kind of two-dimensional contingency table that helps to illustrate how well an algorithm is doing (see Table 2). This table shows the expected occurrences of a class in each row and the actual examples of that class in each column.

Table 2   Disinformation matrix

		Actual classes	
		Injury	*No-injury*
Predicted classes	Injury	TP	FP
	No-injury	FN	TN

Note: True negative, false negative, false positive, and true positive are all acronyms that
       mean different things, even though they all mean different things.

The confusion matrix (Table 1) (Figure 16, green leaves) is used to produce several measures for measuring the efficacy of machine learning classifiers: The ratio of actual positive results to the sum of all positive results is a good indicator of sensitivity or accuracy. Total positives comprise both genuine and false positives (FP). This situation has a positive class consisting of injured persons. Examining the percentage of cases that are accurately classified to each class as a function of the total number of examples in that class allows us to gauge the performance of a classifier. One measure of a test's accuracy is its specificity, which is also called recall. It can be calculated by adding up all the samples that should have been positive $(1 + FN)$ or by dividing the total number of samples by the number of positive ones. An indicator of the number of instances correctly discovered inside a given class is the recall of a classifier. Two times the mean of the recall and precision divided by the total of the two is the F1-score. There is a harmonious relationship between sensitivity (recall) and specificity (accuracy). One measure of performance that accounts for all possible levels of classification is the area under the curve (AUC). It is the likelihood that, when presented with two independent examples, a model will prioritise the positive event. One can get AUC scores between half a point and one.

A higher level of model accuracy is associated with a larger AUC. Injury prediction models cannot be adequately evaluated by using classifier performance as the metric for

the mean of the two classes. This is due to the fact that the dataset is highly unbalanced, which causes the no-injury class to impact the average value significantly. Examining the differences between the two sets of data can shed light on the model's accuracy. The quality of the injury class measures utilised to train machine learning algorithms determines how good the systems are. By determining the average and standard deviation of the metric values discovered for each fold, a cross-validation method may be used to evaluate the models' validity and variability. The standard deviation of the training data has a greater impact on machine learning models, making them less stable. In an evolutionary setting, however, we can evaluate the efficacy of the machine learning algorithms on a daily or weekly basis. On top of that, you may assess its efficacy in the evolutionary scenario by looking at the classifiers' cumulative performance, which starts from the first day of prediction and goes onwards.

5 Conclusions

This study presents research that shows how DT technology might improve sports injury prediction and performance management. By leveraging a multi-stage DT development framework – integrating biomechanical modelling, dimensionality reduction, and advanced machine learning algorithms – the study has successfully demonstrated how digital replicas of athletes can be used to forecast injury risks with significant accuracy. Among the algorithms evaluated, XGBoost outperformed others, achieving a prediction accuracy of 78% and identifying critical biomechanical indicators such as hamstring force, muscle stiffness, and fibre velocity as top predictors. Furthermore, the incorporation of explainable AI (XAI) techniques allowed for greater transparency in model decisions, offering actionable insights for coaches, clinicians, and sport scientists. The biomechanical simulations performed through musculoskeletal modelling also validated the physical plausibility of the predicted injury regions, aligning with real-world physiological behaviour. The credibility and possibility of the system for preventing injuries and improving performance in real-time are highlighted by the fact that the simulated outputs and actual anatomical stress responses are converging. With the use of both conventional and evolutionary data sampling techniques, the study also proved that the DT framework could adapt to different situations.

This highlights the versatility and robustness of the framework, making it suitable for deployment in dynamic sports environments where athlete conditions frequently change. In the future, the model's predictive strength and contextual relevance could be even better by incorporating multi-modal data sources, like psychological and environmental aspects. Real-time deployment through wearable technologies and cloud-based analytics platforms would also support continuous athlete monitoring, opening avenues for personalised and proactive sports management. Overall, the study lays a strong foundation for applying DT technology in sports, marking a significant step toward intelligent, data-driven athlete care and injury prevention.

Declarations

All authors declare that they have no conflicts of interest.

References

Adolfo, C.M.S., Chizari, H., Win, T.Y. and Al-Majeed, S. (2021) 'Sample reduction for physiological data analysis using principal component analysis in artificial neural network', *Applied Sciences*, Vol. 11, No. 17, p.8240.

Alderson, J. and Johnson, W. (2016) 'The personalised digital athlete: an evolving vision for the capture, modelling and simulation of on-field athletic performance', *ISBS – Conference Proceedings Archive*, 6 November [online] https://ojs.ub.uni-konstanz.de/cpa/article/view/7099.

Barricelli, B.R., Casiraghi, E., Gliozzo, J., Petrini, A. and Valtolina, S. (2020) 'Human digital twin for fitness management', *IEEE Access*, Vol. 8, pp.26637–26664, https://doi.org/10.1109/ACCESS.2020.2971576.

Batty, M. (2024) 'Digital twins in city planning', *Nature Computational Science*, Vol. 4, No. 3, pp.192–199.

Bettega, K. and Düggelin, W. (2023) 'Innovating the digital product development toolchain with digital twins at Helbling Technik AG', *Journal of Information Technology Teaching Cases*, https://doi.org/10.1177/20438869231182450.

Calderón-Díaz, M. and Silvestre Aguirre, R. (2024) 'Explainable machine learning techniques to predict muscle injuries in professional soccer players through biomechanical analysis', *Sensors*, 25 December, Vol. 24, No. 1, p.119.

Celik, Y., Powell, D. and Woo, W.L. (2020) 'A feasibility study towards instrumentation of the sport concussion assessment tool (iSCAT)', in *Proc. 2020 42nd Annual Int. Conf. IEEE Eng. in Med. & Biology Society (EMBC)*, pp.4624–4627, https://doi.org/10.1109/EMBC44109.2020.9175656.

Ferrari, A. and Willcox, K. (2024) 'Digital twins in mechanical and aerospace engineering', *Nature Computational Science*, Vol. 4, No. 3, pp.178–183.

Gabrišová, I. et al. (2025) 'the adoption of modern sports technologies from professional settings to everyday life', *Administrative Sciences*, Vol. 15, No. 7, p.249.

Glebova, E., Book, R., Su, Y., Perić, M. and Heller, J. (2023) 'Sports venue digital twin technology from a spectator virtual visiting perspective', *Frontiers in Sports and Active Living*, Vol. 5, p.1289140, https://doi.org/10.3389/fspor.2023.1289140.

Hliš, T. and Fister, I. (2024) 'Digital twins in sport: concepts, taxonomies, challenges and practical potentials', *Expert Systems with Applications*, Vol. 258, p.125104, https://doi.org/10.1016/j.eswa.2024.125104.

Jiang, Y., Yin, S. and Li, K. (2021) 'Industrial applications of digital twins', *Philosophical Transactions of the Royal Society A*, Vol. 379, No. 2207, p.20200360, https://doi.org/10.1098/rsta.2020.0360.

Jones, D., Snider, C. and Nassehi, A. (2020) 'Characterising the digital twin: a systematic literature review', *CIRP Journal of Manufacturing Science and Technology*, 1 May, Vol. 29, pp.36–52.

Laubenbacher, R. and Mehrad, B. (2024) 'Digital twins in medicine', *Nature Computational Science*, Vol. 4, No. 3, pp.184–191.

Mănescu, D.C. (2025) 'Big data analytics framework for decision-making in sports performance optimization', *Data*, Vol. 10, p.116, https://doi.org/10.3390/data10070116.

Mendes-Neves, T. and Meireles, L. (2023) *A Survey of Advanced Computer Vision Techniques for Sports*, arXiv, https://arxiv.org/abs/2301.07583.

Miehe, R. and Waltersmann, A. (2021) 'Sustainable production and the role of digital twins-basic reflections and perspectives', *Journal of Advanced Manufacturing & Processing*, Vol. 3, No. 2, p.e10078, https://doi.org/10.1002/amp2.10078.

Mundt, M. (2023) 'Bridging the lab-to-field gap using machine learning in sports biomechanics', *Sports Biomechanics*, Vol. 22, No. 4, pp.1–20.

Powell, D., Stuart, S. and Godfrey, A. (2021) 'Sports-related concussion: an emerging era in digital sports technology', *NPJ Digital Medicine*, Vol. 4, No. 1, p.164, https://doi.org/10.1038/s41746-021-00538-w.

Rico-González, M. and Los Arcos, A. (2020) 'Accuracy and reliability of local positioning systems for measuring sport movement patterns in stadium-scale: a systematic review', *Applied Sciences*, 29 August, Vol. 10, No. 17, p.5994.

Rocha, C., Lobo, J. and Parente, M. (2025) 'Biomechanical modeling and simulation of the knee joint: integration of AnyBody and Abaqus', *Biomechanics*, Vol. 5, No. 3, p.57.

Rossi, A. and Pappalardo, L. (2022) 'A narrative review for a machine learning application in sports: an example based on injury forecasting in soccer', *Sports*, 24 December, Vol. 10, No. 1, p.5.

Stuart, S. et al. (2020) 'Analysis of free-living mobility in people with mild traumatic brain injury and healthy controls: quality over quantity', *Journal of Neurotrauma*, Vol. 37, pp.139–145, https://doi.org/10.1089/neu.2019.6450.

Tokas, D. and Kumar, H. (2024) 'Embracing the 'digital athlete': how ai and digital technology can improve sports injury management: a thematic review', *International Journal of Physical Education, Sports and Health*, Vol. 11, No. 4, pp.560–568.