



OPEN

Leveraging explainable machine learning to identify gait biomechanical parameters associated with anterior cruciate ligament injury

Christos Kokkotis^{1,2✉}, Serafeim Moustakidis³, Themistoklis Tsatalas², Charis Ntakolia^{4,5}, Georgios Chalatsis⁶, Stylianos Konstadakos⁷, Michael E. Hantes⁶, Giannis Giakas² & Dimitrios Tsaopoulos¹

Anterior cruciate ligament (ACL) deficient and reconstructed knees display altered biomechanics during gait. Identifying significant gait changes is important for understanding normal and ACL function and is typically performed by statistical approaches. This paper focuses on the development of an explainable machine learning (ML) empowered methodology to: (i) identify important gait kinematic, kinetic parameters and quantify their contribution in the diagnosis of ACL injury and (ii) investigate the differences in sagittal plane kinematics and kinetics of the gait cycle between ACL deficient, ACL reconstructed and healthy individuals. For this aim, an extensive experimental setup was designed in which three-dimensional ground reaction forces and sagittal plane kinematic as well as kinetic parameters were collected from 151 subjects. The effectiveness of the proposed methodology was evaluated using a comparative analysis with eight well-known classifiers. Support Vector Machines were proved to be the best performing model (accuracy of 94.95%) on a group of 21 selected biomechanical parameters. Neural Networks accomplished the second best performance (92.89%). A state-of-the-art explainability analysis based on SHapley Additive exPlanations (SHAP) and conventional statistical analysis were then employed to quantify the contribution of the input biomechanical parameters in the diagnosis of ACL injury. Features, that would have been neglected by the traditional statistical analysis, were identified as contributing parameters having significant impact on the ML model's output for ACL injury during gait.

Anterior cruciate ligament (ACL) tear is a frequent knee injury occurring in young active individuals during sport activities like basketball, football, ski and volleyball^{1,2}. The primary function of the ACL is to confine excessive posterior translation and external rotation of the femur relatively to the tibia against forces that act on the joint during gait and other activities³⁻⁶. As a result, an ACL deficient knee presents significant reflect on joint stability and biomechanics⁷⁻⁹. Studies utilising three-dimensional (3D) motion analysis have shown altered joint motion in ACL deficient knees during daily activities, such as walking, ascending and descending stairs or jumping¹⁰⁻¹². This deviation causes a shift on the contact area and magnitude of shear forces at the knee joint which can lead to the initiation of osteoarthritis¹³⁻¹⁶.

ACL reconstruction (ACLR) aims to lessen these changes in knee biomechanics. Annually 130.000 ACL reconstruction surgeries are performed in United States¹⁷. Although ACLR provides an improvement in knee stability and kinematics it is still questionable if the results are equal to pre-injury standards^{18,19}. As it was observed

¹Institute for Bio-Economy & Agri-Technology, Center for Research and Technology Hellas, 38333 Vólos, Greece. ²TEFAA, Department of Physical Education & Sport Science, University of Thessaly, 42100 Trikala, Greece. ³AIDEAS OÜ, 10117 Tallinn, Estonia. ⁴Hellenic National Center of COVID-19 Impact on Youth, University Mental Health Research Institute, 11527 Athens, Greece. ⁵School of Naval Architecture and Marine Engineering, National Technical University of Athens, 15772 Athens, Greece. ⁶Department of Orthopaedic Surgery and Musculoskeletal Trauma, Faculty of Medicine, School of Health Sciences, University General Hospital of Larissa, 41110 Larissa, Greece. ⁷Animus, 41500 Larissa, Greece. ✉email: chkokkotis@gmail.com

in several studies, increase or decrease in peak external knee-adduction moment, peak internal-rotation angle, increased medial contact force and decreased knee flexion angles were related to knee-joint cartilage loading and degeneration^{13,20–22}. Reductions in peak knee-flexion angle and external knee-flexion moment during the loading phase of gait have been reported at 6 to 60 months after ACLR^{23–25}.

Machine learning (ML) is an artificial intelligence (AI) analytic tool that constructs algorithms to identify patterns and characteristics contained within datasets. The goal is to train and validate prediction algorithms to achieve a desired result²⁶. Musculoskeletal-specific models have already been developed to identify and classify fractures and predict functional outcomes after primary total knee arthroplasty (TKA)²⁷. In 2017, Olczak et al. used deep learning techniques based on medical imaging to examine the feasibility of using AI to identify fractures in skeletal radiographs²⁸. In another study, Kunze et al. based on partially modifiable risk factors developed ML algorithms to predict dissatisfaction after TKA²⁹. Recent studies with individual-level datasets of gait analyses from kinetic skeletal tracking and advanced MR imaging (MRI) techniques focused on the determination of early progression of knee osteoarthritis (KOA)³⁰. Moustakidis et al. proposed a novel fuzzy decision tree-based support vector machine (SVM) classifier by using 3-D ground reaction force (GRF) measurements to investigate KOA severity and to distinguish between asymptotic and osteoarthritis knee gait patterns³¹. Furthermore, Padoia et al. performed ML multidimensional data analysis by using MR imaging and biomechanical data³². They demonstrated that the analysis potentially indicates that cartilage composition may be an imaging biomarker for early KOA.

Machine learning approaches have been also used in studies to identify ACL injury based on MRI and biomechanical data or ACLR gait patterns with the aid of motion sensors. In 2017, Mazlan et al. proposed an ACL injury diagnosis system by using ACL injury MRI (normal, partial and crucial ACL) and SVM algorithm³³. In another study, Chang et al. used MRI and deep learning techniques for the detection of complete ACL tear and achieved 96% test set accuracy³⁴. Furthermore, Christian et al. used gait kinematics and ML techniques (SVM) to develop a pattern recognition system for diagnosis and evaluation of therapeutic treatment effect³⁵. In another study, Zeng et al. proposed an approach for detection of the presence of ACL injury using kinematic features and neural networks³⁶. Moreover, Todesco et al. proposed an ML approach for the identification of ACL gait patterns based on motion sensors data for on the field activities in rugby players³⁷.

Despite the relatively large number of ML studies on the field of ACL, the reported trained ML models are treated as black boxes. The lack of transparency and explainability of the models result to poor understanding of their inner workings and the rationale behind their decision-making mechanism. This paper focuses on the development of an explainable ML-empowered methodology to identify important biomechanical parameters associated with ACL injury. The aims of this study are: (i) to estimate the feature importance in the classification process and examine how much each of the features contributed to the final ML decisions and (ii) to investigate differences in sagittal plane kinematics and kinetics of the gait cycle between different patient groups based on a novel approach that combines explainable ML and statistical analytics. To achieve these goals, an extensive experimental setup was designed including biomechanical data collection, a thorough comparative analysis with seven well-known classifiers and a state-of-the-art explainability analysis.

Results

Comparative analysis. The proposed ML pipeline was initially applied on the three-class problem in which the patient groups CON, ACLD and ACLR are considered as separate classes. The proposed FS technique was executed on the pre-processed version of the 3-class dataset ranking the available features with respect to their relevance. The ML models were trained on feature subsets of increasing dimensionality (with a step of 1) and the testing classification accuracies were finally calculated until the full feature set has been tested. The classification results are given below.

Figure 1 demonstrates the accuracy testing performance (%) of the competing ML models with respect to the number of selected features on the 3-class problem. The majority of the ML models had an upward trend in the whole feature dimensionality range, followed by steady testing performance in most of the cases. Specifically, the SVM model showed an upward trend with respect to the first selected features, with a maximum of 94.95% (which was the overall best performance achieved). The second-best accuracy (92.89%) was achieved by the NN model, which presented a non-steadily increasing performance with fluctuations for more than 15 selected features. LR, DTs, RF, XGboost and KNN models also showed an upward trend with moderate accuracies ranging from 68.18% up to 90.40%. In contrast with the other models, Naïve Bayes failed in this task, recording low accuracy testing performances (in the range of 44.44–59.09%).

Table 1 summarizes the results of XGboost, Random Forest, Decision Trees, Naive Bayes, SVM, KNN, Logistic regression and NN on the three-class problem. The number of selected biomechanical parameters, on which the highest accuracy was obtained, lies within the range of 19–23 for the majority of the ML models (in seven out of the eight), whereas the overall maximum performance was achieved by SVM on a group of twenty-one selected (21) biomechanical parameters. Naive Bayes selected less features (14) leading to low accuracy (57.58%). Furthermore, the second-highest accuracy was achieved by NN (92.89%), whereas lower accuracies were obtained by RF, DTs and XGboost (less than 74.75%). Apart from being the most accurate overall, the SVM model recorded the best performance in all the metrics, namely precision (92.16–96.72%), recall (92.19–97.62%) and f1-score (93.07–96.47%).

Explainability to quantify the features' contribution. In this section, we interpret the contribution of the biomechanical parameters in shaping the AI model's output. To cope with this, we used explainability analysis on the best performing ML model (SVM). Initially, we performed a global investigation on the 3-class problem to quantify the overall features' contribution to the problem. Then, we performed explainability analy-

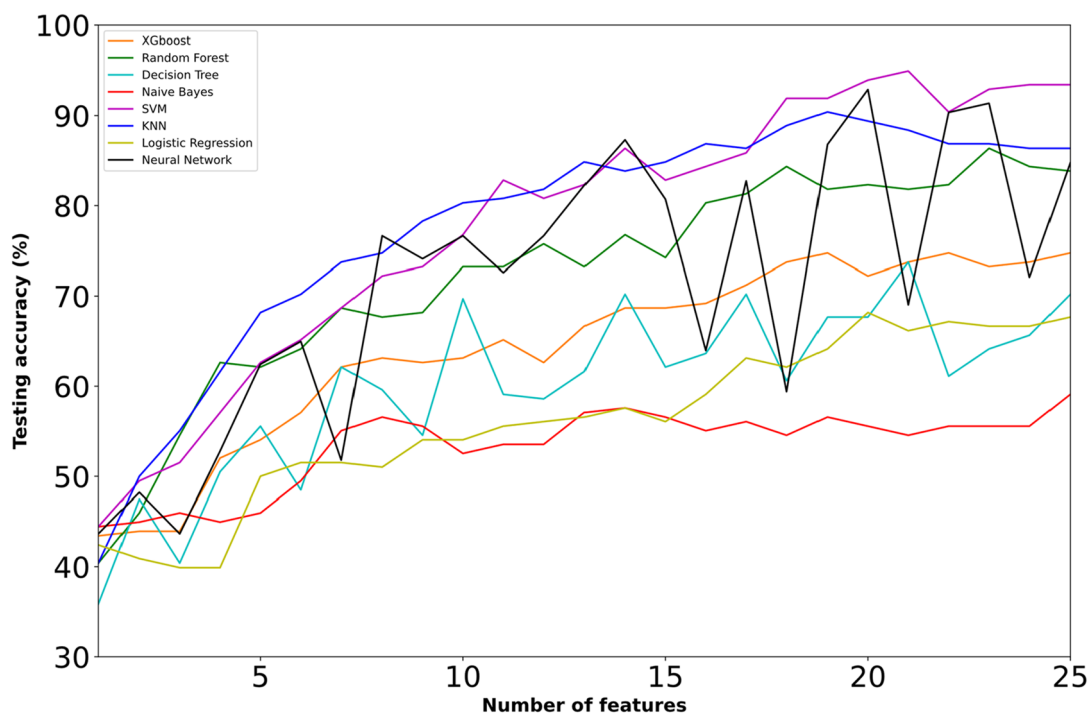


Figure 1. Learning curves with testing accuracy scores for different ML models trained on feature subsets of increasing dimensionality in the 3-class problem (referring to both ACL deficient and ACL reconstructed patients).

Models	Accuracy	Classes	Precision	Recall	F1-Score	Num. of features
XGBoost	74.75	CON	70.80	95.24	81.22	19
		ACLD	81.48	44.00	57.14	
		ACLR	79.31	71.88	75.41	
Random Forest	86.36	CON	80.00	95.24	86.96	23
		ACLD	90.00	72.00	80.00	
		ACLR	94.83	85.94	90.16	
Decision Trees	73.74	CON	76.67	82.14	79.31	21
		ACLD	72.09	62.00	66.67	
		ACLR	70.77	71.88	71.32	
Naive Bayes	57.58	CON	65.69	79.76	72.04	14
		ACLD	40.91	54.00	46.55	
		ACLR	66.67	31.25	42.55	
SVM	94.95	CON	95.35	97.62	96.47	21
		ACLD	92.16	94.00	93.07	
		ACLR	96.72	92.19	94.40	
KNN	90.40	CON	85.26	96.43	90.50	19
		ACLD	95.12	78.00	85.71	
		ACLR	95.16	92.19	93.65	
Logistic Regression	68.18	CON	70.64	91.67	79.79	20
		ACLD	57.50	46.00	51.11	
		ACLR	71.43	54.69	61.95	
NN	92.89	CON	96.30	92.86	94.55	20
		ACLD	90.57	96.00	93.20	
		ACLR	90.48	90.48	90.48	

Table 1. Best testing accuracies (%) achieved for ML models in 3-class problem along with precision, recall, f1-score and the optimum number of features.

sis on each one of the three trained binary (one-versus-one) SVM models that constitute the 3-class problem. Specifically, we applied SHAP analysis into the following three problems: (i) control group versus ACLD group (local problem 1), (ii) control group versus ACLR group (local problem 2), and (iii) ACLD group versus ACLR group (local problem 3).

Figure 2a visualises the impact of the feature across all classes and the features were sorted by the sum of their SHAP values magnitudes across all instances. In this approach K2, H4, A3, GRF4, GRF7, K1, A4 and GRF6 were the parameters that affected the model output with mean SHAP values higher than 0.3.

Figure 2b depicts the mean absolute value of the SHAP values which represents the SHAP global feature importance for local problem 1 (CON versus the ACLD). It should be noted that the features H4, K7, GRF3, H1, H2 were the most important variables that significantly affected the prediction output. It is also observed that the contribution of H4 is 0.3 while the second-best parameter (K7) and all the remaining ones are below 0.18. From the above, H4 significantly contributes to the separation between the CON group and the ACLD group.

Figure 2c depicts the mean absolute value of the SHAP values for local problem 2 that focuses on the discrimination of the CON and ACLR groups. Features K2, GRF7, H4, GRF4 and K1 were the most important variables that significantly affected the prediction output for the certain groups. Specifically, K2 records a much higher mean absolute value (higher than 0.35) compared to the rest of the features (that exhibit values less than 0.23).

The most important variables that significantly affected the prediction output in the local problem 3 (ACLD group versus ACLR group) were K2, H3, K7, A5 and A2, as shown in Fig. 2d. Similarly to local problem 2, parameter K2 is again the most important separation factor between individuals from the ACLD group and the ACLR group.

Statistical analysis to identify significant differences between selected features. Statistical comparisons were also performed to identify whether there exist statistically significant differences between the classes considered for the most important features as they have been highlighted by the explainability analysis. Table 2 demonstrates the results of the one-way ANOVA tests, which performed to quantify these differences at the global level (with all three classes considered). As observed, there were statistically significant differences between the group means for all the comparisons.

Then we performed statistical comparisons at the local level putting emphasis on the following tasks: (i) ACL diagnosis and (ii) rehabilitation after surgery. Initially, we run independent t-test analysis between the CON and the ACLD groups for the first eight significant biomechanical parameters, which were indicated by the explainability analysis of the specific binary problem (local problem 1). Subsequently, we employed independent t-test analysis between the control and the ACL-reconstructed groups on the same parameters to identify which of them were modified and/or restored to their normal state (control level) as a measure of evaluating the postoperative progress.

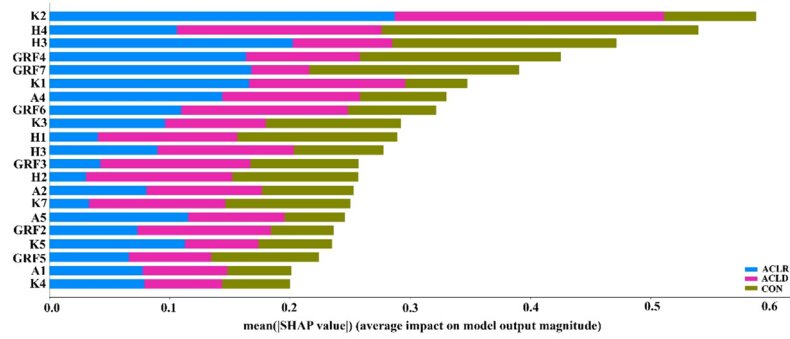
Table 3 summarizes the results of the statistical analysis at the local level. The following remarks can be drawn from Table 5: (i) Significant differences were observed between CON and ACLD for half of the features considered, specifically the first three (H4, K7 and GRF3) along with GRF4; (ii) Four of the parameters (H1, H2, GRF6 and GRF5) that were considered important by the explainability analysis had no significant changes between CON and ACLD groups.

Discussion

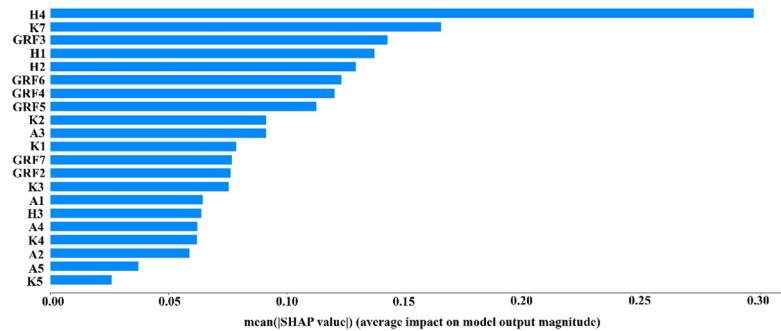
This paper focuses on the development of a novel approach, which combines an explainable ML-empowered methodology and statistical analysis, for identifying important parameters associated with ACL injury. The problem has been coped as a three-class classification task where the participants of the study were divided into three groups (CON, ACLD and ACLR group). In addition to the classification part, the main contributions of this paper are: (i) to investigate how much each of the features contributed to the final ML decisions, (ii) to estimate the feature importance in the classification process and (iii) to investigate differences in three dimensional GRFs, sagittal plane kinematics and kinetics of the gait cycle for the CON, ACLD and ACLR groups.

Being effective in problems with strong dependencies between features, the Relieff algorithm was applied to serve as a FS technique and thus reduce the dimensionality of the initial feature space. Eight ML models were employed to perform the 3-class classification task on the reduced feature space where accuracies up to 94.95% were achieved. Specifically, the SVM model had the best performance and it showed an upward trend with respect to the first selected features, with a maximum of 94.95% at 21 features (which was the overall best performance achieved). Furthermore, the SVM model achieved rates from 92.16% up to 97.62% in each class for the metrics precision, recall and f1-score.

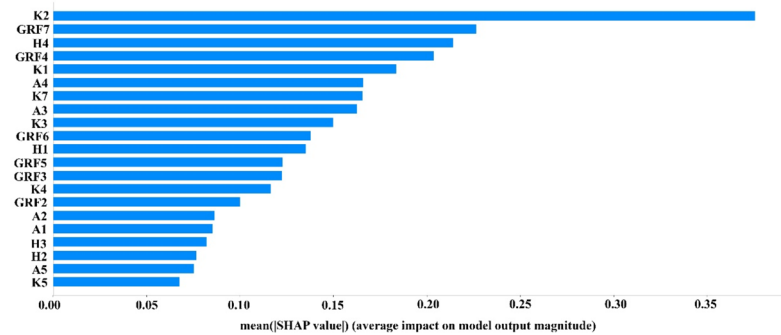
Having selected the most accurate ML model, this study attempted to uncover the rationale behind the decision-making mechanism of the trained model and therefore provide an alternative and a more holistic approach of quantifying the contribution of the input biomechanical parameters in the classification process. Specifically, explainability analysis was applied on the best performing ML model (SVM) and a global investigation was initially performed on the 3-class problem to quantify the overall features' contribution to the problem. As observed K2, H4, A3, GRF4, GRF7, K1, A4 and GRF6 were the most important biomechanical parameters that affected the model output. In order to estimate the feature importance separately, we also performed explainability analysis on each one of the three trained binary (one-versus-one) SVM models that constitute the 3-class problem. Specifically, we applied SHAP analysis into the following three problems: (i) CON group versus ACLD group (local problem 1), (ii) CON group versus ACLR group (local problem 2), and (iii) ACLD group versus ACLR (local problem 3). As observed, in the local problem 1 the main biomechanical parameters were H4, K7 and GRF3. Furthermore, K2, GRF7 and H4 have the main contribution in local problem 2. In addition, from the third local problem K2, H3 and K7 have occurred as the most important biomechanical parameters. Previous studies have



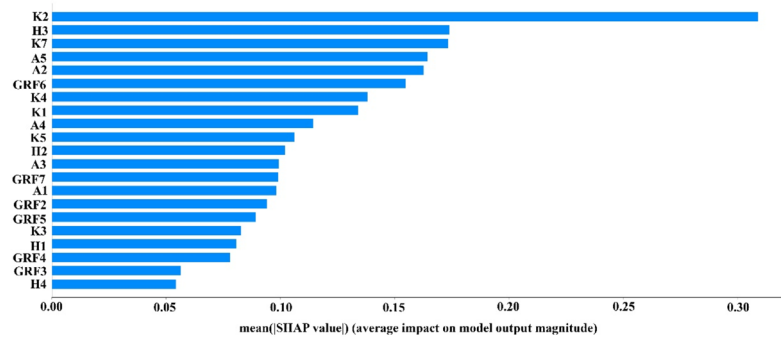
(a)



(b)



(c)



(d)

Figure 2. (a) Average feature impact magnitude for all instances in the 3-class problem; (b) Features’ impact on SVM model output for local problem 1. This figure shows the average impact magnitude for all instances in the task of differentiating the control group vs pre-surgery group; (c) Average feature impact magnitude for all instances in the local problem 2 (control versus ACLR); (d) Average feature impact magnitude for all instances for local problem 3 (pre-surgery group versus post-surgery group).

Features	Statistical comparison	CON	ACLD	ACLR
		Mean (std)	Mean (std)	Mean (std)
K2	P=0.004	8.59±3.81	8.35±4.89	9.72±4.70
H4	P=0.007	37.93±4.93	36.26±6.42	37.03±5.53
A3	P=0.000	13.81±6.52	15.91±7.43	16.88±7.25
GRF4	P=0.000	19.61±4.30	16.94±5.17	16.76±4.84
GRF7	P=0.000	5.19±1.65	5.81±2.03	6.18±2.92
K1	P=0.001	21.62±5.88	19.73±6.56	19.88±6.27
A4	P=0.011	0.18±0.08	0.19±0.08	0.20±0.09
GRF6	P=0.004	5.69±1.43	6.01±2.21	6.36±2.97

Table 2. Statistical comparison at the global level.

Features*	CON vs ACLD	CON vs ACLR
H4	P=0.002	P=0.057
K7	P=0.000	P=0.001
GRF3	P=0.000	P=0.090
H1	P=0.288	P=0.792
H2	P=0.723	P=0.326
GRF6	P=0.061	P=0.001
GRF4	P=0.000	P=0.000
GRF5	P=0.721	P=0.147

Table 3. Statistical analysis at the local level for ACL diagnosis and rehabilitation.

observed altered gait biomechanics in the ACL deficient and ACL reconstructed patients compared to healthy individuals^{10,11}. These findings may indicate that the employed rehabilitation protocols fail to restore normal walking biomechanics, resulting in aberrant movement patterns. Several of the most important biomechanical parameters of ACL injury diagnosis highlighted by the global as well as the local explainability analysis used in our study coincide with the biomechanical outcomes reported in the literature to be related to altered gait patterns following ACLR. For example, maximum knee extension moment during stance phase (K7) significantly affected the prediction output in both of the aforementioned local problems examined in our study. K7 has been extensively investigated following ACLR and it has been consistently identified as a biomechanical parameter that is decreased following surgery and it is associated with poorer knee function in ACLR patients compared to healthy individuals^{11,38}. Additionally, minimum knee flexion angle during stance phase (K2) which had the most important contribution in local problem 2 and a significant one in local problem 3 has been reported to differentiate gait patterns between ACLR and healthy individuals up to 48 weeks post-surgery³⁸.

Besides explainability analysis, conventional statistical analysis was further performed to determine whether there exist significant differences between the three groups of our study for the aforementioned selected biomechanical parameters. As it was observed, in most of the cases the outcomes of the explainability and statistical analyses coincide. However, no significant differences were identified for many of those important parameters as shown in the case of local problem 1 (ACL diagnosis) in which H1, H2, GRF6 and GRF5 were identified as important by SHAP whereas their distributions had no significant differences between CON and ACLD. This finding implies that the proposed explainable ML methodology goes beyond the way that traditional statistics work. Features, that would have been neglected by the traditional statistical analysis, are highlighted as contributing parameters that have a significant impact on the ML model's output when they are combined with other statistically important ones. Moreover, as a measure of evaluating the postoperative progress, we performed statistical analysis for the local problem 1 and local problem 3 on the same parameters to identify which of them were modified and/or restored to their normal state (control level) after the surgery. Two (H4 and GRF3) of the three most important parameters (H4, GRF3, the third parameter) were restored to their initial state after the surgery having no significant differences in the comparison between CON and ACLR groups. This means that these two biomechanical parameters (H4 and GRF3) were initially modified after the ACL injury and they were subsequently restored to their initial state after the surgery.

The clinical significance of our novel approach discussed in this work, which is based on a combination of an explainable ML-empowered methodology and statistical analysis to identify biomechanical parameters during walking associated with ACL injury, should be considered with caution. This can be attributed to the fact that even though gait biomechanics are altered following ACLR, few biomechanical parameters demonstrate consistent results across studies and various tasks¹⁰. Factors such as, differences in the ACLR techniques (e.g. graft type), individual coping strategies among participants during walking, variations in employed rehabilitation

Characteristics	ACLD	ACLR	CON
Gender	31 males and 13 females	40 males and 14 females	34 males and 19 females
Height	175.3 ± 0.86 cm	177.6 ± 0.80 cm	174.1 ± 0.98 cm
Weight	77.38 ± 14.91 kg	76.37 ± 14.35 kg	72.23 ± 15.81 kg

Table 4. Subjects' characteristics.

protocols and gender differences may affect gait biomechanics alterations following ACLR as well as their clinical interpretation^{10,11,39}.

Explainability via SHAP or other similar tools is a crucial enabler allowing humans to better comprehend the decisions generated by black box models. However, SHAP is limited to simple explanations mainly quantifying the impact of individual features to the models' output⁴⁰. Thus, the inner workings of the trained models and the way that the features are combined to reach the final decision remain hidden. Future work includes the combined use of graphical modelling with well-known explainability tools with the goal of identifying the relationships between features and the possible direct and indirect effect of features to the models' output. Such graphically-given explanations would enhance our understanding of the real rationale behind the decision-making mechanism of ML-empowered models acting on the tasks of ACL diagnosis and rehabilitation.

Summary

An explainable ML-empowered methodology was designed, implemented and tested in this paper to identify important biomechanical parameters associated with ACL injury. The proposed extensive experimental setup included gait biomechanical data, a thorough comparative analysis with seven well-known classifiers and a state-of-the-art explainability analysis. According to the findings of the comparative analysis, a 94.95% classification accuracy was achieved by SVM on a group of twenty-one biomechanical parameters. The nature of the selected parameters along with their impact on the prediction outcome (via SHAP) were discussed to uncover the rationale behind the decision-making mechanism of the trained model and therefore provide an alternative and a more holistic approach of quantifying the contribution of the input parameters in the diagnosis of ACL injury. Statistical analysis was further performed to determine whether there exist significant differences between ACL deficient, ACL reconstructed and healthy individuals for the aforementioned parameters. Understanding the contribution of gait biomechanics is a valuable tool for creating more powerful and non-invasive prognostic tools in the hands of physicians, that will point out abnormal gait patterns in patients after ACLR to modify the rehabilitation protocol and avoid the development of osteoarthritis.

Methods

Participants. A total of 151 subjects volunteered to participate in this study. Three different groups were defined: (i) ACL-deficient prior to surgery (ACLD), (ii) ACL-reconstructed (ACLR) and (iii) control (CON) group. Inclusion criteria were as follows: (1) all subjects could be of either gender between 18–50 years old and were moderately active, participating in regular activity at least two times per week; (2) The ACLD subjects had suffered a unilateral ACL injury confirmed by an orthopedic surgeon as well as via magnetic resonance imaging and participated in the present study an average of 30 days after injury, but before surgery; (3) The ACLR subjects had a unilateral ACL reconstruction and participated in the present study at least 6 months post-surgery; (4) Individuals with different graft types (i.e., hamstring tendon and patellar tendon grafts) were included in the ACLR group; (5) Both ACLD and ACLR subjects had a healthy contralateral knee and reported no other history of serious lower limb injury, such as bilateral ACL injury or injury to the meniscus, posterior cruciate ligament, medial or lateral collateral ligament in either knee; (6) had been cleared to resume their physical activity at the time of the measurement. The CON subjects were matched for age, gender, and physical activity status and had no history of ACL injury and neurologic disorder or other lower extremity injuries within 12 months prior to participating in the study. Prior to participation, all subjects signed a consent form, and all procedures were approved by the University of Thessaly ethics committee (approval code: 1660). The subjects' characteristics are presented in detail in Table 4.

Testing procedure and data collection. Upon entering the gait laboratory, the subjects received instructions regarding the testing procedure and were familiarized with the walking task. ACLD and ACLR subjects completed the subjective Knee injury and Osteoarthritis Outcome Score (KOOS) evaluation form, which is considered a reliable measure of 5 outcomes, including activities of daily living, sport and recreation, pain, and knee-related quality of life⁴¹. Anthropometric measurements were recorded, and 20 spherical retroreflective markers were positioned bilaterally on anatomic landmarks and specific locations of the pelvis and lower limbs according to the marker set described in the literature^{42,43}. Subsequently, the subjects walked barefoot along the 10 m laboratory walkway within ± 5% of their individual self-selected walking speed (SWS). SWS was measured during familiarization using infrared timing gates located in the middle of the walkway and was maintained throughout data collection via a metronome. Trials were performed until at least 5 complete gait cycles were recorded with each foot (left and right side) landing on the force platform. A trial was considered valid if the foot of the side being tested made a clean contact with the force platform located in the middle of the walkway and the walking speed was within ± 5% of the individual SWS. Kinematic data were collected using 10 optoelectronic

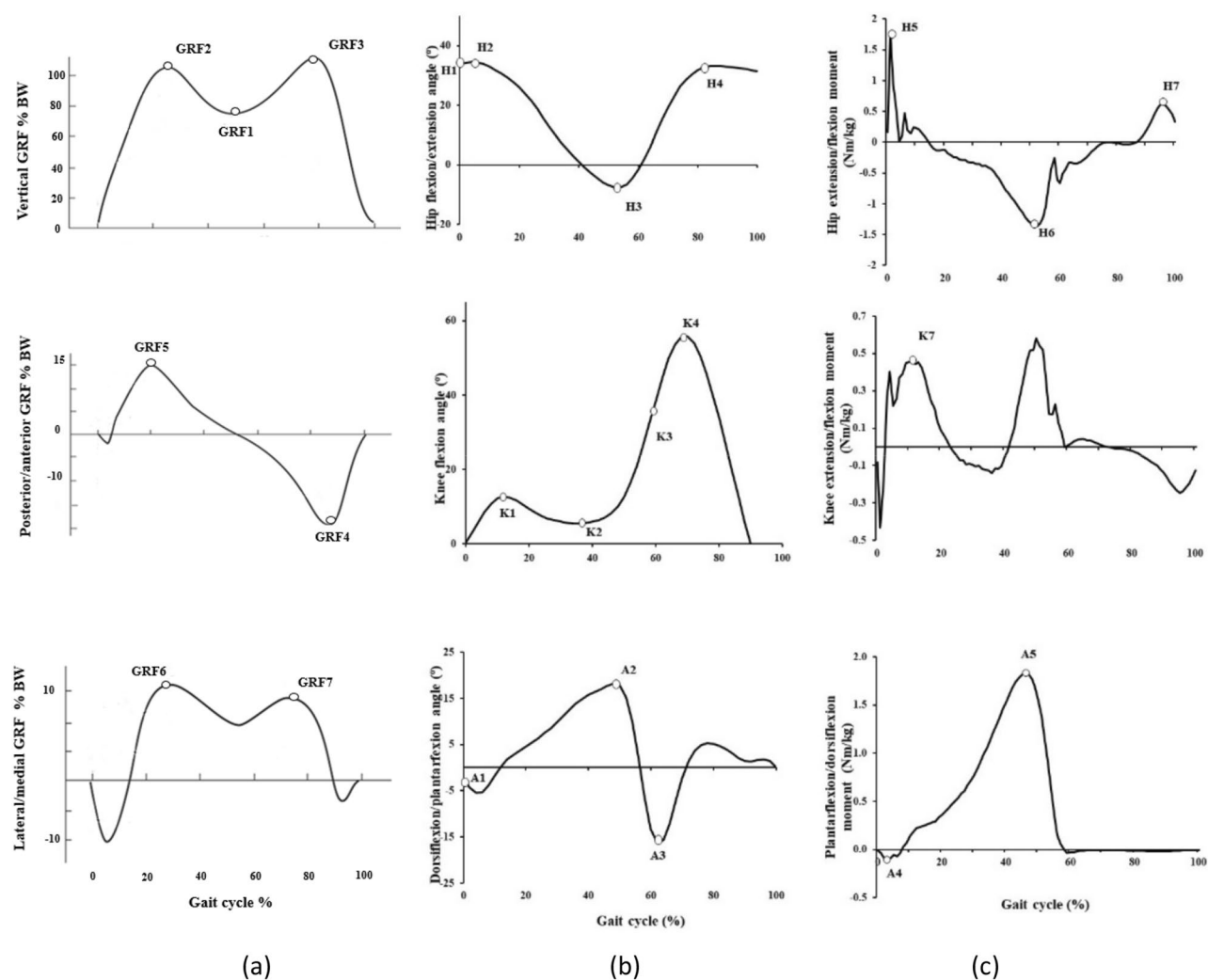


Figure 3. Three dimensional GRFs (a), sagittal plane kinematic (b) and kinetic (c) variables of interest during walking. Ankle dorsiflexion, knee flexion, hip flexion, anterior and medial GRFs, as well as internal ankle plantar flexion, knee extension and hip extension moments were all defined as positive.

cameras (Vicon T-series, Oxford, UK) at 100 Hz and kinetic data were collected at 1000 Hz via a force platform (Bertec 4060–10, OH) embedded in the floor synchronized with the kinematic data.

Data analysis. The symmetrical center of rotation estimation (SCoRE)⁴⁴ and the symmetrical axes of rotation approach (SARA)⁴⁵ were applied to optimize the calculation of the hip joint center and knee joint flexion axis, respectively. The initial contact and toe-off events of stance phase were determined from the vertical GRF (20 N threshold) and the subsequent ipsilateral initial contact was determined from motion data using the Vicon Nexus software. Kinematic and GRF data were lowpass filtered with a 4th order Butterworth filter at 10 and 40 Hz, respectively. The SCoRE and SARA approach were combined with the Plug-In-Gait model to generate motion features according to the Nexus Advanced Gait Workflow (Nexus 2.10, Vicon Metrics Group Ltd, Oxford, UK). Inverse dynamics were used combining inertia properties of the segments as well as sagittal plane kinematic and GRF data to calculate net joint moments of the lower limbs during the gait cycle. GRFs were expressed as a percentage of body weight, while net joint moments were expressed as internal moments and were normalized to body mass. Selected gait variables were extracted for each trial of each subject. A total of 155 trials were analyzed for ACLD group, 204 trials for ACLR group and 298 trials for CON group, respectively. The three-dimensional GRFs, sagittal plane kinematic and kinetic variables of interest are presented in Fig. 3 and Table 5. Data were analyzed from the ACLD/ACLR subjects' involved limb and for the control subjects, this was randomly assigned.

Machine learning workflow. In order to identify knee kinematics associated with ACL injury, we designed, implemented and tested a multi-stage ML pipeline. Data were normalised to [0, 1] to build a common basis for the feature selection (FS) and the ML estimators. To rank our biomechanical parameters a well-established FS technique was applied. Relief algorithm⁴⁶ is a supervised learning algorithm and it is such effective in problems

Variables	Description
GRF1	Local minimum vertical GRF during support (% BW)
GRF2	First vertical GRF peak (% BW)
GRF3	Second vertical GRF peak (% BW)
GRF4	Anterior (propulsive) GRF peak (% BW)
GRF5	Posterior (braking) GRF peak (% BW)
GRF6	First medial GRF peak (% BW)
GRF7	Second medial GRF peak (% BW)
H1	Hip flexion angle at initial contact (°)
H2	Maximum hip flexion angle during stance phase (°)
H3	Maximum hip extension angle during stance phase (°)
H4	Maximum hip flexion angle during swing phase (°)
H5	Maximum hip extension moment during stance phase (Nm/kg)
H6	Maximum hip flexion moment during stance phase (Nm/kg)
H7	Maximum hip extension moment during swing phase (Nm/kg)
K1	Peak knee flexion angle during stance phase (°)
K2	Minimum knee flexion angle during stance phase (°)
K3	Knee flexion angle at foot off (°)
K4	Maximum knee flexion angle during swing phase (°)
K5	Knee flexion angle at first maximum knee extension moment during stance phase (°)
K6	Knee flexion angle at first vertical ground reaction force peak (°)
K7	First maximum knee extension moment during stance phase (Nm/kg)
A1	Ankle angle at initial contact (°)
A2	Maximum dorsi-flexion angle during stance phase (°)
A3	Maximum plantar-flexion angle over the entire gait cycle (°)
A4	Maximum dorsiflexion moment during stance phase (Nm/kg)
A5	Maximum plantarflexion moment during stance phase (Nm/kg)

Table 5. Evaluated parameters of the gait cycle for vertical and horizontal GRFs and sagittal plane kinematics and kinetics.

where strong dependencies between features are observed. Various well-known ML classifiers were evaluated for their suitability. Hyperparameter selection was implemented to avoid bias error, overfitting and optimize the performance of our ML models. Specifically, we used XGboost algorithm⁴⁷ and Random Forest (RF)⁴⁸, which are ensemble learning algorithms and they are used due to their fast execution speed and increased model performance. Decision trees (DTs) were also evaluated providing a more interpretable decision-making mechanism⁴⁹. Furthermore, we tested Naïve Bayes algorithm⁵⁰, which is based on applying Bayes' theorem and this method can be extremely fast. Being effective in high-dimensional spaces, SVM algorithms were also included in our experimental analysis⁵¹. Moreover, Logistic Regression (LR)⁵² and the K-Nearest Neighbor (KNN) algorithm⁵³ were tested. LR was employed to set the baseline performance obtained by a linear model and KNN was selected due to its ability to deal with the overfitting problems that appear in high-dimensional spaces. Neural networks (NN) were also tested because they can handle complex data using a human brain inspired structure that mimics the way that biological neurons communicate to each other⁵⁴. Different activation functions were tested (including tanh, sigmoid and ReLU) for their suitability as part of the hyperparameter optimisation phase.

For the evaluation of the proposed classifiers, a stochastic 70–30% random data split was applied to generate the training and testing subsets, respectively⁵⁵. Specifically, the learning was performed on the stratified version of the training sets and the final performance was estimated on the accuracy testing sets. Furthermore, the performance of the classifiers was also evaluated in terms of the recall (or sensitivity), f1-score and precision as additional evaluation criteria⁵⁶.

In this paper, we also: (i) investigated how much each of the features contributed to the final decision and (ii) estimated the feature importance. In order to achieve this, we used SHapley Additive exPlanations (SHAP) which are based on Shapley Values of game theory^{57,58}. SHAP offers the ability to interpret ML models, which are often treated as black boxes. In this paper, we employed SHAP to rank features in terms of their impact on the final ML outputs and to build a mini explainer model. This enhances our understanding of the internal decision-making rationale of the trained AI models especially with respect to the mechanism with which selected biomechanical parameters are combined to produce decisions on ACL diagnosis and rehabilitation.

Statistical analysis. One-way analysis of variance (ANOVA) was used to investigate differences in sagittal plane kinematics and kinetics of gait cycle for the CON, ACLD and ACLR groups⁵⁹. Furthermore, independent sample t-tests were employed to compare the first eight significant biomechanical parameters between the CON and the ACLD groups, which were indicated by the explainability analysis. On the same parameters, independ-

ent sample t-tests were also employed to evaluate the postoperative progress⁶⁰. The significance level in our statistical comparisons was set at $P < 0.05$.

Institutional review board statement. The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the University of Thessaly (protocol code 1660 and date of approval 03/06/2020).

Informed consent. Informed consent was obtained from all subjects involved in the study.

Data availability

The dataset generated during and/or analysed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

Received: 13 September 2021; Accepted: 11 April 2022

Published online: 22 April 2022

References

1. Prodromos, C. C., Han, Y., Rogowski, J., Joyce, B. & Shi, K. A meta-analysis of the incidence of anterior cruciate ligament tears as a function of gender, sport, and a knee injury-reduction regimen. *Arthroscopy* **23**, 1320–1325.e1326. <https://doi.org/10.1016/j.arthro.2007.07.003> (2007).
2. Moses, B., Orchard, J. & Orchard, J. Systematic review: Annual incidence of ACL injury and surgery in various populations. *Res. Sports Med.* **20**, 157–179. <https://doi.org/10.1080/15438627.2012.680633> (2012).
3. Kanamori, A. *et al.* The effect of axial tibial torque on the function of the anterior cruciate ligament: A biomechanical study of a simulated pivot shift test. *Arthroscopy* **18**, 394–398. <https://doi.org/10.1053/jars.2002.30638> (2002).
4. Zantop, T., Herbort, M., Raschke, M. J., Fu, F. H. & Petersen, W. The role of the anteromedial and posterolateral bundles of the anterior cruciate ligament in anterior tibial translation and internal rotation. *Am. J. Sports Med.* **35**, 223–227. <https://doi.org/10.1177/0363546506294571> (2007).
5. Hanzlíková, I. *et al.* The effect of proprioceptive knee bracing on knee stability during three different sport related movement tasks in healthy subjects and the implications to the management of Anterior Cruciate Ligament (ACL) injuries. *Gait Posture* **48**, 165–170 (2016).
6. Rezende, F. C. *et al.* Does combined intra- and extra-articular ACL reconstruction improve function and stability? A meta-analysis. *Clin. Orthop. Relat. Res.* **473**, 2609–2618 (2015).
7. Andriacchi, T. P. & Dyrby, C. O. Interactions between kinematics and loading during walking for the normal and ACL deficient knee. *J Biomech* **38**, 293–298. <https://doi.org/10.1016/j.jbiomech.2004.02.010> (2005).
8. Georgoulis, A. D., Papadonikolakis, A., Papageorgiou, C. D., Mitsou, A. & Stergiou, N. Three-dimensional tibiofemoral kinematics of the anterior cruciate ligament-deficient and reconstructed knee during walking. *Am. J. Sports Med.* **31**, 75–79. <https://doi.org/10.1177/03635465030310012401> (2003).
9. Tsarouhas, A. *et al.* Three-dimensional kinematic and kinetic analysis of knee rotational stability after single- and double-bundle anterior cruciate ligament reconstruction. *Arthroscopy* **26**, 885–893. <https://doi.org/10.1016/j.arthro.2009.11.016> (2010).
10. Moore, J. M. *et al.* Lower limb biomechanics before and after anterior cruciate ligament reconstruction: A systematic review. *J. Biomech.* **106**, 109828. <https://doi.org/10.1016/j.jbiomech.2020.109828> (2020).
11. Hart, H. F. *et al.* Knee kinematics and joint moments during gait following anterior cruciate ligament reconstruction: A systematic review and meta-analysis. *Br. J. Sports Med.* **50**, 597–612. <https://doi.org/10.1136/bjsports-2015-094797> (2016).
12. DeFrate, L. E. *et al.* The 6 degrees of freedom kinematics of the knee after anterior cruciate ligament deficiency: An in vivo imaging analysis. *Am. J. Sports Med.* **34**, 1240–1246. <https://doi.org/10.1177/0363546506287299> (2006).
13. Andriacchi, T. P., Koo, S. & Scanlan, S. F. Gait mechanics influence healthy cartilage morphology and osteoarthritis of the knee. *J. Bone Jt. Surg.* **91**(Suppl 1), 95–101. <https://doi.org/10.2106/jbjs.h.01408> (2009).
14. Andriacchi, T. P., Briant, P. L., Beville, S. L. & Koo, S. Rotational changes at the knee after ACL injury cause cartilage thinning. *Clin. Orthop. Relat. Res.* **442**, 39–44. <https://doi.org/10.1097/01.blo.0000197079.26600.09> (2006).
15. Chaudhari, A. M., Briant, P. L., Beville, S. L., Koo, S. & Andriacchi, T. P. Knee kinematics, cartilage morphology, and osteoarthritis after ACL injury. *Med. Sci. Sports Exerc.* **40**, 215–222. <https://doi.org/10.1249/mss.0b013e31815cbb0e> (2008).
16. Butler, R. J., Minick, K. I., Ferber, R. & Underwood, F. Gait mechanics after ACL reconstruction: Implications for the early onset of knee osteoarthritis. *Br. J. Sports Med.* **43**, 366–370. <https://doi.org/10.1136/bjsm.2008.052522> (2009).
17. Mall, N. A. *et al.* Incidence and trends of anterior cruciate ligament reconstruction in the United States. *Am. J. Sports Med.* **42**, 2363–2370. <https://doi.org/10.1177/0363546514542796> (2014).
18. Slater, L. V., Hart, J. M., Kelly, A. R. & Kuenze, C. M. Progressive Changes in Walking Kinematics and Kinetics After Anterior Cruciate Ligament Injury and Reconstruction: A Review and Meta-Analysis. *J. Athl. Train.* **52**, 847–860. <https://doi.org/10.4085/1062-6050-52.6.06> (2017).
19. Papannagari, R. *et al.* In vivo kinematics of the knee after anterior cruciate ligament reconstruction: A clinical and functional evaluation. *Am. J. Sports Med.* **34**, 2006–2012. <https://doi.org/10.1177/0363546506290403> (2006).
20. Andriacchi, T. P. & Mündermann, A. The role of ambulatory mechanics in the initiation and progression of knee osteoarthritis. *Curr. Opin. Rheumatol.* **18**, 514–518. <https://doi.org/10.1097/01.bor.0000240365.16842.4e> (2006).
21. Hurwitz, D. E. *et al.* Knee pain and joint loading in subjects with osteoarthritis of the knee. *J. Orthop. Res.* **18**, 572–579. <https://doi.org/10.1002/jor.1100180409> (2000).
22. Sharma, L. *et al.* Knee adduction moment, serum hyaluronan level, and disease severity in medial tibiofemoral osteoarthritis. *Arthritis Rheum.* **41**, 1233–1240. [https://doi.org/10.1002/1529-0131\(199807\)41:7%3c1233::aid-art14%3e3.0.co;2-l](https://doi.org/10.1002/1529-0131(199807)41:7%3c1233::aid-art14%3e3.0.co;2-l) (1998).
23. Di Stasi, S. L., Logerstedt, D., Gardinier, E. S. & Snyder-Mackler, L. Gait patterns differ between ACL-reconstructed athletes who pass return-to-sport criteria and those who fail. *Am. J. Sports Med.* **41**, 1310–1318. <https://doi.org/10.1177/0363546513482718> (2013).
24. Timoney, J. M. *et al.* Return of normal gait patterns after anterior cruciate ligament reconstruction. *Am. J. Sports Med.* **21**, 887–889. <https://doi.org/10.1177/036354659302100623> (1993).
25. Shin, C. S., Chaudhari, A. M., Dyrby, C. O. & Andriacchi, T. P. Influence of patellar ligament insertion angle on quadriceps usage during walking in anterior cruciate ligament reconstructed subjects. *J. Orthop. Res.* **27**, 730–735. <https://doi.org/10.1002/jor.20806> (2009).
26. Bayliss, L. & Jones, L. D. The role of artificial intelligence and machine learning in predicting orthopaedic outcomes. *Bone Jt. J.* **101-B**, 1476–1478. <https://doi.org/10.1302/0301-620x.101b12.bjj-2019-0850.r1> (2019).

27. Sanchez-Santos, M. T. *et al.* Development and validation of a clinical prediction model for patient-reported pain and function after primary total knee replacement surgery. *Sci. Rep.* **8**, 3381–3381. <https://doi.org/10.1038/s41598-018-21714-1> (2018).
28. Olczak, J. *et al.* Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop.* **88**, 581–586. <https://doi.org/10.1080/17453674.2017.1344459> (2017).
29. Kunze, K. N., Polce, E. M., Sadauskas, A. J. & Levine, B. R. Development of machine learning algorithms to predict patient dissatisfaction after primary total knee arthroplasty. *J. Arthroplasty* **35**, 3117–3122. <https://doi.org/10.1016/j.arth.2020.05.061> (2020).
30. Brisson, N. M., Gatti, A. A., Damm, P., Duda, G. N. & Maly, M. R. Association of machine learning based predictions of medial knee contact force with cartilage loss over 2.5 years in knee osteoarthritis. *Arthr. Rheumatol.* **73**, 1638–1645. <https://doi.org/10.1002/art.41735> (2021).
31. Moustakidis, S. P., Theocharis, J. B. & Giakas, G. Feature selection based on a fuzzy complementary criterion: Application to gait recognition using ground reaction forces. *Comput. Methods Biomech. Biomed. Eng.* **15**, 627–644. <https://doi.org/10.1080/10255842.2011.554408> (2012).
32. Pedoia, V. *et al.* MRI and biomechanics multidimensional data analysis reveals R2–R1rho as an early predictor of cartilage lesion progression in knee osteoarthritis. *J. Magn. Reson. Imaging* **47**, 78–90. <https://doi.org/10.1002/jmri.25750> (2018).
33. Mazlan, S., Ayob, M. Z. & Bakti, Z. Anterior cruciate ligament (ACL) injury classification system using support vector machine (SVM). In *2017 International Conference on Engineering Technology and Technopreneurship (ICE2T)*, 1–5 (2017).
34. Chang, P. D., Wong, T. T. & Rasiej, M. J. Deep learning for detection of complete anterior cruciate ligament tear. *J. Digit. Imaging* **32**, 980–986. <https://doi.org/10.1007/s10278-019-00193-4> (2019).
35. Christian, J. *et al.* Computer aided analysis of gait patterns in patients with acute anterior cruciate ligament injury. *Clin. Biomech.* **33**, 55–60. <https://doi.org/10.1016/j.clinbiomech.2016.02.008> (2016).
36. Zeng, W., Ismail, S. A. & Pappas, E. Detecting the presence of anterior cruciate ligament injury based on gait dynamics disparity and neural networks. *Artif. Intell. Rev.* **53**, 3153–3176. <https://doi.org/10.1007/s10462-019-09758-9> (2020).
37. Tedesco, S. *et al.* Motion sensors-based machine learning approach for the identification of anterior cruciate ligament gait patterns in on-the-field activities in Rugby players. *Sensors* **20**, 3029 (2020).
38. Knoll, Z., Kocsis, L. & Kiss, R. M. Gait patterns before and after anterior cruciate ligament reconstruction. *Knee Surg. Sports Traumatol. Arthrosc.* **12**, 7–14. <https://doi.org/10.1007/s00167-003-0440-1> (2004).
39. Asaeda, M. *et al.* Gender differences in the restoration of knee joint biomechanics during gait after anterior cruciate ligament reconstruction. *Knee* **24**, 280–288 (2017).
40. Dikopoulou, Z., Moustakidis, S. & Karlsson, P. GLIME: A new graphical methodology for interpretable model-agnostic explanations. [arXiv:2107.09927](https://arxiv.org/abs/2107.09927) (2021). <https://ui.adsabs.harvard.edu/abs/2021arXiv210709927D>.
41. Roos, E. M. & Lohmander, L. S. The Knee injury and Osteoarthritis Outcome Score (KOOS): From joint injury to osteoarthritis. *Health Qual. Life Outcomes* **1**, 1–8 (2003).
42. Tsatalas, T. *et al.* Altered drop jump landing biomechanics following eccentric exercise-induced muscle damage. *Sports* **9**, 24 (2021).
43. Tsatalas, T. *et al.* Walking kinematics and kinetics following eccentric exercise-induced muscle damage. *J. Electromyogr. Kinesiol.* **23**, 1229–1236. <https://doi.org/10.1016/j.jelekin.2013.04.008> (2013).
44. Ehrig, R. M., Taylor, W. R., Duda, G. N. & Heller, M. O. A survey of formal methods for determining the centre of rotation of ball joints. *J. Biomech.* **39**, 2798–2809. <https://doi.org/10.1016/j.jbiomech.2005.10.002> (2006).
45. Ehrig, R. M., Taylor, W. R., Duda, G. N. & Heller, M. O. A survey of formal methods for determining functional joint axes. *J. Biomech.* **40**, 2150–2157. <https://doi.org/10.1016/j.jbiomech.2006.10.026> (2007).
46. Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S. & Moore, J. H. Relief-based feature selection: Introduction and review. *J. Biomed. Inform.* **85**, 189–203 (2018).
47. Torlay, L., Perrone-Bertolotti, M., Thomas, E. & Baciu, M. Machine learning–XGBoost analysis of language networks to classify patients with epilepsy. *Brain Informatics* **4**, 159–169 (2017).
48. Svetnik, V. *et al.* Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **43**, 1947–1958 (2003).
49. Podgorelec, V., Kokol, P., Stiglic, B. & Rozman, I. Decision trees: An overview and their use in medicine. *J. Med. Syst.* **26**, 445–463 (2002).
50. Miranda, E., Irwansyah, E., Amelga, A. Y., Maribondang, M. M. & Salim, M. Detection of cardiovascular disease risk's level for adults using naive Bayes classifier. *Healthc. Inform. Res.* **22**, 196–205 (2016).
51. Ma, Y. *et al.* Classification of motor imagery EEG signals with support vector machines and particle swarm optimization. *Comput. Math. Methods Med.* **2016**, 1–8 (2016).
52. Subasi, A. & Ercelebi, E. Classification of EEG signals using neural network and logistic regression. *Comput. Methods Programs Biomed.* **78**, 87–99 (2005).
53. Park, J. & Lee, D. H. Privacy preserving k-nearest neighbor for medical diagnosis in e-health cloud. *J. Health. Eng.* **2018**, 1–11 (2018).
54. Mathan, K., Kumar, P. M., Panchatcharam, P., Manogaran, G. & Varadharajan, R. A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. *Des. Autom. Embed. Syst.* **22**, 225–242 (2018).
55. Kokkotis, C., Moustakidis, S., Giakas, G. & Tsaopoulos, D. Identification of risk factors and machine learning-based prediction models for knee osteoarthritis patients. *Appl. Sci.* **10**, 6797 (2020).
56. Ghosh, M. & Sanyal, G. An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning. *J. Big Data* **5**, 1–25 (2018).
57. Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S. & Mohammadian, A. K. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accid. Anal. Prev.* **136**, 105405 (2020).
58. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. in *Advances in neural information processing systems*. 4765–4774 (2017).
59. Bewick, V., Cheek, L. & Ball, J. Statistics review 9: One-way analysis of variance. *Crit. Care* **8**, 1–7 (2004).
60. Gerald, B. A brief review of independent, dependent and one sample t-test. *Int. J. Appl. Math. Theor. Phys.* **4**, 50–54 (2018).

Acknowledgements

This work was supported by the EC Horizon 2020 project OACTIVE Grant Agreement No. 777159, by the European Union and Greek National Funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH–CREATE–INNOVATE (SafeACL Project, Grant Agreement T1EDK-04234) and by the Postgraduate Program of Study “Military Fitness & Wellbeing”, School of Physical Education, Sports Science, University of Thessaly, Greece.

Author contributions

Design and conduct of the study (C.K., S.M., D.T.); data collection (C.K., T.T., G.H., S.K.); data processing (C.K., T.T., C.N., S.K.); manuscript preparation (C.K., S.M., T.T., C.N., G.H.); manuscript review (M.E., G.G., D.T.).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022