



HAL
open science

How can machine learning help prevent knee injuries of experienced athletes?

Clément Lipps, Julien Frère, Thierry Weissland

► **To cite this version:**

Clément Lipps, Julien Frère, Thierry Weissland. How can machine learning help prevent knee injuries of experienced athletes?. ECSS 2023 - 28th Annual Congress of the European College of Sport Science, INSEP French Institute of Sport, Jul 2023, Paris, France. hal-04252695

HAL Id: hal-04252695

<https://hal.science/hal-04252695v1>

Submitted on 23 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How can machine learning help prevent knee injuries of experienced athletes ?

Clément LIPPS LENE¹, Julien FRERRE², Thierry WEISSLAND¹

¹PMH_DySCo, IMS, UMR 5218, Université de Bordeaux, CNRS, Talence, France, ²MOVE, GIPSA-Lab, UMR 5216, Univ Grenoble Alpes, CNRS, Grenoble INP, Grenoble, France

Introduction: Knee injuries are among the most frequently occurring sports-related injuries [1] and are highly multifactorial. As machine learning (ML) algorithms have the ability to understand complex relationships in large datasets, they seem appropriate to predict at-risk athletes based on a large number of features [2]. While decision trees (DT) is the most used ML techniques due to its interpretability, recent progress in explainable artificial intelligence (XAI) allow us to consider more complex techniques such as gradient boosting (XGB) and deep feedforward neural networks (MLPs) [3]. Thus, we compared the ability of these three artificial intelligence models to recognize the presence of knee injuries sequelae in experienced athletes.

Protocol: Data were collected from a cohort of 96 athletes practicing on average 5.7±2.4 hours a week, 36 of them had a history of knee injury while the 60 others hadn't. Participants answers an informative questionnaire, a psychologic questionnaire (KOOS, SAS) and performed a horizontal force-velocity test on a ballistic ergometer (Fig 1, [4]). The test consisted in maximal horizontal push-off in supine position on a frictionless cart under five resistive condition (0%, 30%, 60%, 90% and 110% of the individual bodyweight). Two force plates [Kistler 9260AA3] and a linear encoder (sampling frequency : 2000 Hz) were used to record ground forces and cart displacement. Data preprocessed are presented in Table 1, Table 2 and Table 3.

Methods: DT, MLP and XGB are trained to identify the presence of knee injury sequelae in one push-off. The training has been done using 1000 cross validations made on athletes. Metrics used to compare algorithms performance are the accuracy, the precision, the recall, the F1-score and the area under the roc curve (AUC).

For each participant, a prediction score has been computed by dividing the number of push offs classified as 'presence of sequelae' by the number of push offs performed.

Results : On one thousand cross validations, the best model to recognize sequelae of knee injury in pushes was the MLP (Table 4). Fig 2 presents the results of MLP for the classification of the retained test set athletes. Table 5 presents the performance metrics for athlete classification for each best trained model. To explain predictions, SHAP value were computed for the best trained MLP. Fig 3 presents the most important risk factors are classified from top to bottom. Men with a low score in KOOS Quality of life, a high BMI and practicing at a high level of competition are more likely to have knee injury sequelae according to the best MLP model.

Hypothesis: XGB and MLPs will outperform DT methods in knee injury sequelae recognition

Feature	Unit
History of knee injury	1 if yes, else 0
Height	cm
Weight	kg
Body mass index	kg.m ⁻²
Age	years
Gender	unitless
Sport practiced	unitless
Consistency of practice	hours/week
Level of practice	unitless

Table 1: The 9 informative variables extracted from the questionnaire

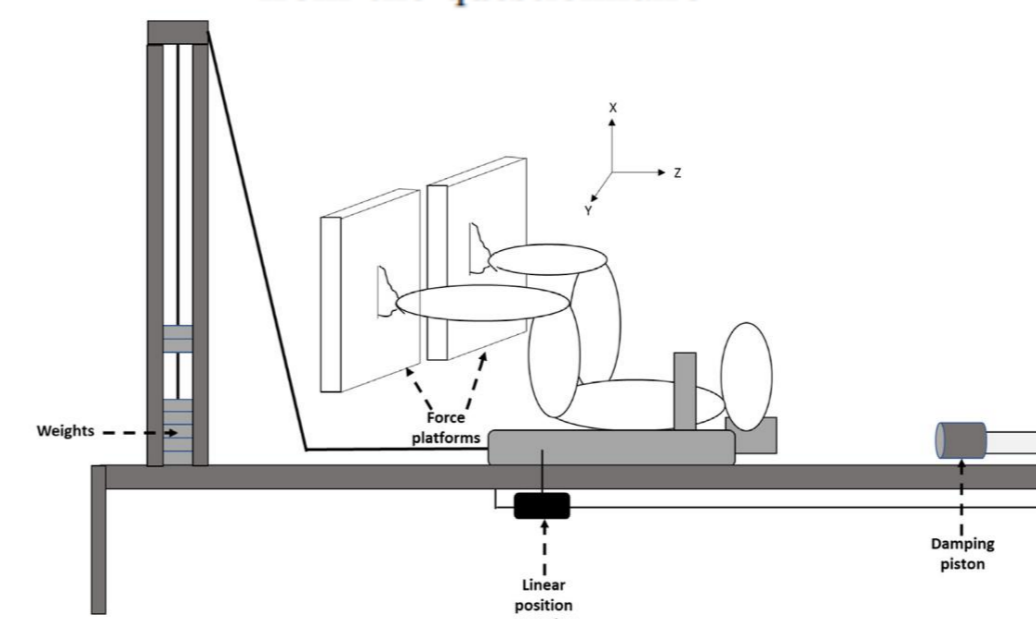


Fig 1 : Set-up on the horizontal force-velocity ergometer [4]

Feature	Unit
SAS Score	unitless
KOOS Symptoms and recreation	unitless
KOOS Pain	unitless
KOOS Sport	unitless
KOOS Quality of life	unitless

Table 2: The 5 scores extracted from the Sport Anxiety Scale (SAS) the Knee injury and the Osteoarthritis Outcome Score (KOOS)

Feature	Unit
Force peak total	N
Force peak L, Force peak R	N
Force mean total	N
Force mean L, Force mean R	N
Velocity mean	m.s ⁻¹
Inertial force mean total	N
Inertial force mean L, Inertial force mean R	N
Push time	s
Inertial push time	s
Push distance	m
Effective force total ratio	unitless
Effective force ratio L, Effective force ratio R	unitless
Time to reach Force peak total	s
Maximal power	W
Force at maximal power	N
Velocity at maximal power	m.s ⁻¹

Table 3: The 21 biomechanical variables extracted from each push under each resistive condition

Modèle	Accuracy	Precision	Recall	F1 score	AUC
DT	0.89	0.82	0.91	0.86	0.89
XGB	0.88	0.90	0.80	0.85	0.87
MLP	0.92	0.87	0.96	0.91	0.92

Table 4: Maximum value of each metric for each algorithm over the thousand cross validation made on our dataset made of 964 rows (total of push of our 96 athletes) and 35 columns (Table 1 + Table 2 + Table 3 + load level)

Modèle	Accuracy	Precision	Recall	F1 score	AUC
DT	0.85	0.78	0.88	0.82	0.85
XGB	0.85	0.86	0.75	0.80	0.83
MLP	0.90	0.80	1.0	0.89	0.92

Table 5: Best performance metrics scores for athletes classification with the best trained model : average of the metrics on every push-off performed by an athlete

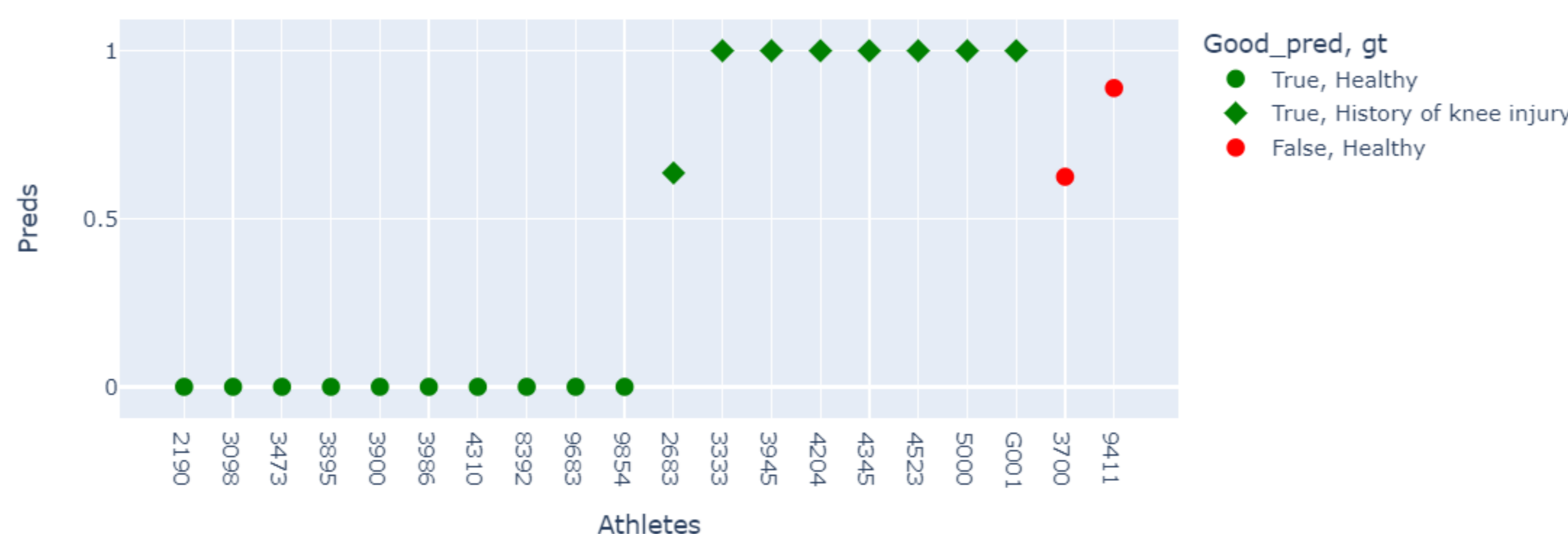


Fig 2 : prediction scores for MLP

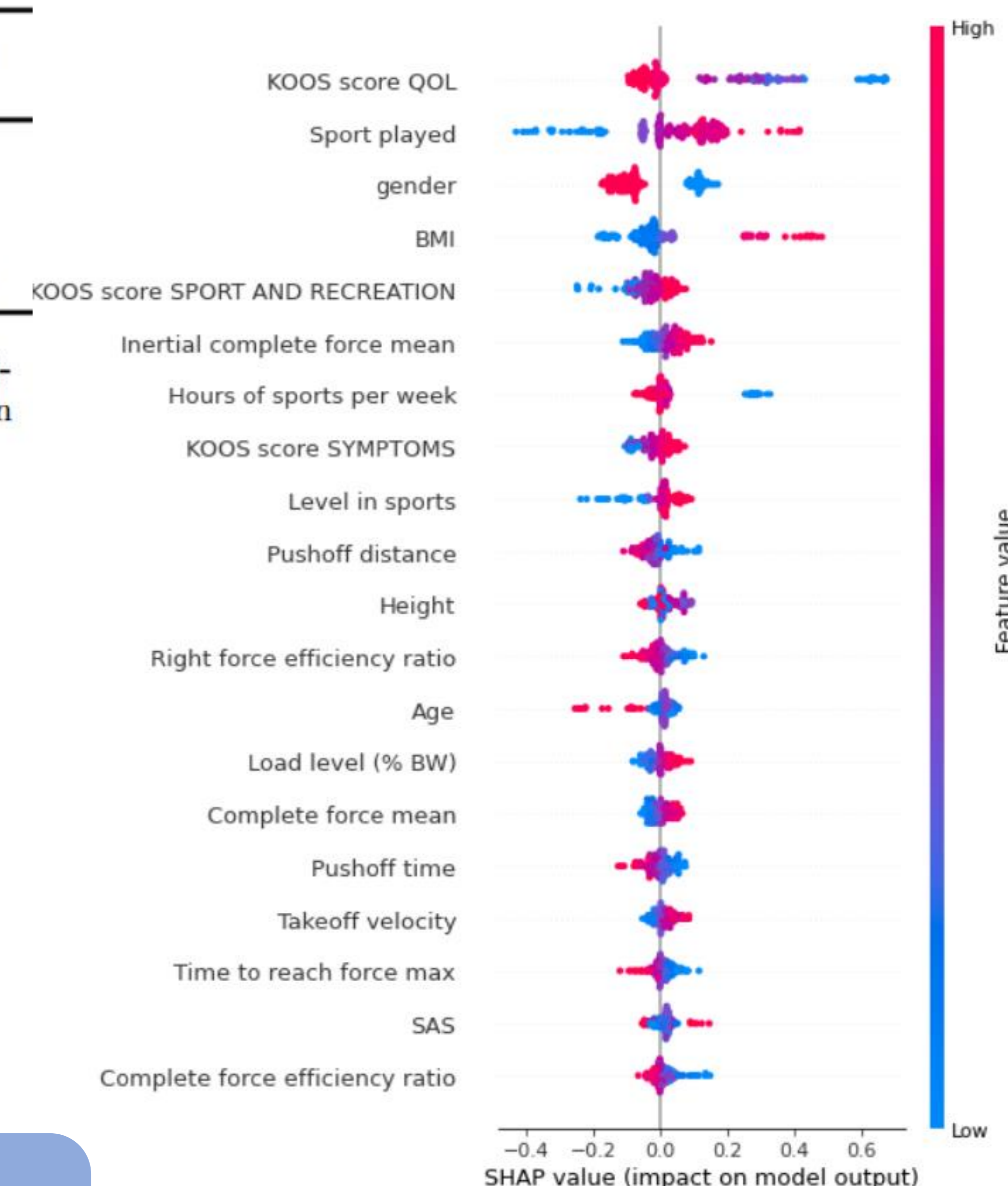


Fig 3 : In general, the most important attributes for the best MLP classification of push-offs and their impact on the prediction. Red points means high feature value, blue points low feature value, points on the right increase the prediction score of knee injury sequelae, points on the left decrease this prediction score

Discussion, Conclusion: This study shows that as expected deep learning based techniques achieves a better result than DT (Table 4) and than previous studies doing the same type of experiment [5]. As a consequence, the most important features for classification are more accurate and can be interpreted as risk factors of knee injuries. Particularly, psychological factors (KOOS scores) and informative attributes (sport played, body mass index) are generally more important than variable describing performances on the test (Fig 3). However, these results must be confirm on a larger cohort like in [5] to have a correct spread of our results.