PREDICTIVE MODELLING OF KNEE ANTERIOR SHEAR FORCES USING INERTIAL MEASUREMENT UNIT DATA AND MACHINE LEARNING

Sean A. Brown[1], Taliah Carlson[1], Lauren Dami[1], and Joshua T. Weinhandl[1]
Kinesiology, Recreation, and Sport Studies, University of Tennessee, Knoxville, TN, USA[1]

This study presents a non-invasive approach using machine learning to predict proximal tibial anterior shear force (ASF), a surrogate for anterior cruciate ligament (ACL) loads from inertial measurement unit (IMU) data, providing a practical alternative to direct force measurement. Employing XGBoost, the research analysed IMU data from drop jump tasks performed by 22 female participants and validated on an additional participant. The model underwent optimization through feature reduction and signal filtering. The results demonstrate an XGBoost model using IMU data to estimate ASF showed improved prediction accuracy after feature reduction and a low-pass filter. The model was able to predict ASF with a root mean squared error of 41.59 ± 13.18N, a mean absolute error of 35.18 ± 11.97N, and an R^2 value of 0.84 ± 0.07.

KEYWORDS: Wearable sensors, XGBoost Regression, Anterior Cruciate Ligament

INTRODUCTION: The anterior cruciate ligament (ACL) is crucial for maintaining knee stability during dynamic movements. Injuries to the ACL are common in athletic populations and can have long-term consequences. Quantifying the forces applied to the ACL, particularly the anterior shear force, is essential for understanding injury mechanisms and for developing preventive measures (Beaulieu et al., 2023). Direct measurement of the forces on the ACL during dynamic activities is not feasible in real-world settings due to the invasive nature required to measure forces transmitted through ligaments. As a result, there is a critical need for non-invasive predictive models that can estimate these forces accurately to aid in injury prevention and rehabilitation strategies (Chappell et al., 2002). Current methods utilize a three-dimensional (3D) motion capture laboratory to collect the movement of participants during specific tasks. However, motion capture labs are expensive to install and require highly trained individuals to operate. Inertial measurement units (IMUs) are inexpensive and easier to utilize to collect data because they do not require highly trained technicians or a dedicated laboratory. Beyond the financial and practical implications, IMUs can be utilized in real world environments, taking the athlete out of the laboratory, and placing them in their real-world environments which may lead to more sport and athlete specific data collection (Fong & Chan, 2010).
The aim of this research was to develop a non-invasive predictive model using machine learning techniques to estimate net anteroposterior joint reaction force, a surrogate for ACL loads (Chappell et al., 2002) from 3D motion capture data. This was accomplished by utilizing data collected via two IMUs during a drop jump movement task. The primary goal was to use the IMU data to predict the proximal tibial anterior shear force (ASF), which was measured and calculated from the 3D motion capture and ground reaction force (GRF) data. This approach intended to leverage the predictive power of machine learning to establish a relationship between the easily obtainable IMU data and the more complex measurements derived from 3D motion capture.

METHODS: 23 healthy, recreationally active females (age: 21.26 ± 1.71 years, height: 1.70 ± 0.05 m, mass: 64.03 ± 8.30 kg) voluntarily participated after providing informed consent. Participants completed 5 drop jump trials with 2 IMUs (500 Hz, IMeasureU, Blue Thunder) affixed to the right shank with elastic straps. One IMU was placed distally over the medial malleolus while the other was placed more proximally just distal of the tibial tuberosity. The IMUs each provide six columns of data, acceleration in three directions (mm/s^2) and angular velocity (degrees/s) as measured by the gyroscopes in three orthogonal planes. The participants simultaneously had 3D marker coordinate data (250 Hz, Vicon Motion Analysis) and 3D GRF data (2000 Hz, BP600600, Advanced Mechanical Technology, Inc.) of the right leg recorded (Weinhandl et al., 2010). IMU data, raw marker data, and GRF data were filtered using a 4^{th} order lowpass Butterworth filter at cutoff frequency of 10 Hz for the IMU data, 6 Hz for the raw marker data, and 50 Hz for the GRF data. ASF were then calculated via inverse dynamic modelling using Visual3D (C-Motion, Germantown, MD, USA). A column-wise linear interpolation was performed to up sample the ASF data to match capture rate of the IMUs.

In this study, data from 22 participants were incorporated into the training and validation dataset, with four of the trials from each participant acting as training data and one trial acting as validation data. One participant's data were reserved exclusively for testing purposes and were not used for the development of the model to provide ASF data from a participant never seen by the model. The model was tested on three trials of the new participant data. A methodological approach was employed to develop and evaluate a machine learning model using extreme gradient boosting (XGBoost), aimed at predicting the ASF data from the IMU data. XGBoost is an advanced gradient boosting framework that enhances predictive modelling by sequentially building decision trees to minimize errors, incorporating regularization to prevent overfitting, and leveraging parallel processing for speed. Its built-in cross-validation feature further ensures model reliability and accuracy by facilitating the assessment of its performance across different data subsets. The data analysis began with the standardization of both input features (IMU data) and the target variable (ASF). This standardization involved adjusting the data to have a mean of zero and a standard deviation of one. Following the initial data preparation, a comprehensive hyperparameter tuning was conducted using a grid search strategy, where a combination of different hyperparameters of the XGBoost algorithm were systematically evaluated. The best combination of parameters was determined through cross-validation using a 10-fold approach and evaluating the model's performance based on the negative mean squared error.

Once the optimal hyperparameters were identified, the model was retrained and evaluated on a separate test dataset. The predictions were adjusted back to the original scale of the target variable for a meaningful evaluation. The evaluation metrics included the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), the coefficient of determination (R^2) score, and a Deming regression. Deming regression, used in method comparison studies where both methods have error, measures agreement between them. A slope near 1 and an intercept close to zero indicate minimal bias and proportional relationship between the methods.

Feature selection was conducted for the XGBoost model, prioritizing the top five features which collectively accounted for 76.9% of the variance. This strategy was adopted to enhance model performance by focusing on the most predictive features, thereby reducing overfitting and eliminating noise and redundancy. The model was retrained and evaluated with a reduced set of features, followed by another evaluation of the same metrics. Then, a 4^{th} order low-pass Butterworth filter with a cutoff frequency of 15 Hz was applied to the model's predictions. This signal processing technique was employed to smooth the predictions and potentially enhance the model's performance by reducing noise in the data. Finally, the performance of the model with the filtered predictions was evaluated on three trials of a test dataset on a new participant never seen by the model, again using RMSE, MAE, and R^2 score as key metrics.

RESULTS: The efficacy of the model was assessed using three distinct configurations: the full feature model, a feature-reduced model, and a filtered model post feature reduction. During feature reduction, the XGBoost model underwent a refinement process, reducing the number of features from 12 to 5.

The full feature XGBoost model yielded a RMSE of 64.08 ± 12.53N (~15% of the average maximum of 408N), a MAE of 51.76 ± 11.74N, and a R^2 of 0.63 ± 0.08 (Table 1). The Deming regression analysis produced a slope of 0.89±0.05 and an intercept of 45.80 ± 8.09N, indicating a proportional bias where predicted values tended to be consistently lower than the actual ASF values. Refinement of the model through feature reduction significantly improved its predictive accuracy, with a reduced RMSE of 48.08±10.82N and MAE of 38.47 ± 10.29 N, and an improved R^2 of 0.79 ± 0.06 (Table 1). The Deming regression slope improved to 1.01±0.06 with an intercept of 14.04 ± 3.55N, reflecting an improved agreement between predicted and actual ASF values. The application of a low-pass filter to the predictions of the feature-reduced model further enhanced performance, achieving the lowest RMSE of 41.59 ± 13.18N and MAE of

35.18 ± 11.97N with an R^2 value increased to 0.84 ± 0.07. The Deming regression slope slightly increased to 1.10 ± 0.06, and the intercept moved closer to zero with a value of -9.64 ± 5.83N (Table 1), indicating a negligible systematic bias and suggesting that the model's predictions closely mirrored the actual ASF values, with a slight tendency to overestimate the ASF when values were low.

Table 1: Model Evaluation Metrics on Test Data for Three Jumps: Mean ± STD

Model	RMSE (N)	MAE (N)	R^2	Deming Regression Slope	Deming Regression Intercept (N)
XGBoost	64.08 ±12.53	51.76 ±11.74	0.63 ±0.08	0.89 ±0.05	45.80 ±8.09
XGBoost with Feature Reduction	48.08 ±10.82	38.47 ±10.29	0.79 ±0.06	1.01 ±0.06	14.04 ±3.55
XGBoost with Feature Reduction and Filtering	41.59 ±13.18	35.18 ±11.97	0.84 ±0.07	1.10 ±0.06	-9.64 ±5.83

Figures 1 and 2 show the evolution of our XGBoost model's performance of predicting the validation data across three stages for one representative trial from the test data: using all features, after feature selection, and following signal filtering. The initial model with all features demonstrates a reasonable approximation of the actual ASF values, but with noticeable variability, particularly during the initial peak (Fig. 2). Post feature selection, the model's predictions appear more aligned with actual values, reflecting the elimination of less informative variables. The final iteration, incorporating a low-pass Butterworth filter, presents the smoothest predictions, closely mirroring the actual ASF trend throughout the activity sequence.

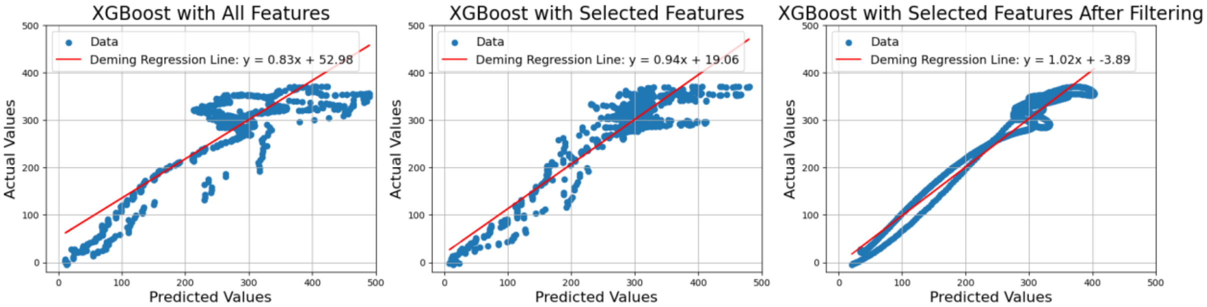

Figure 1: XGBoost model ASF predictions on one representative trial from the test data: full features (left), selected features (middle), and post-filtering (right), with Deming regression lines indicating prediction accuracy and bias.

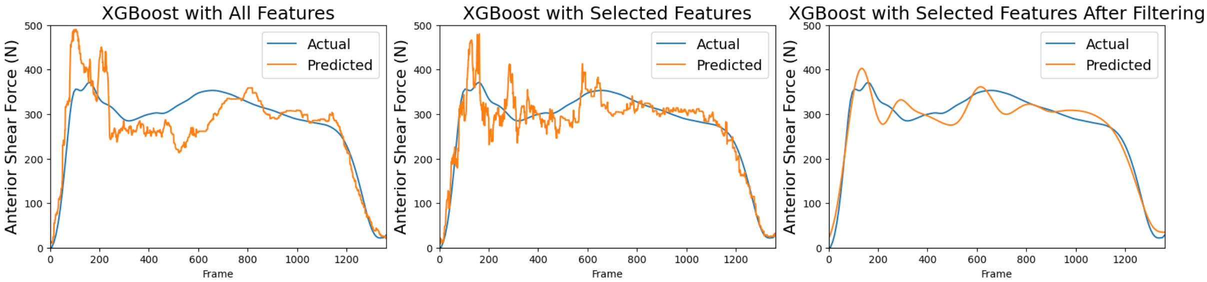

Figure 2: Time-series comparison of actual vs. predicted ASF by XGBoost on one representative trial from the test data: Full features (left), selected features (centre), and post-filtering (right).

DISCUSSION: This investigation into the predictive modelling of ASF via machine learning techniques has yielded promising results. The utilization of IMUs for the estimation of dynamic knee forces offers a practical and non-invasive alternative to conventional motion capture systems, aligning with the growing demand for accessible and real-world applicable biomechanical analyses.

A recent study (Stetter et al., 2019) developed an artificial neural network (ANN) to estimate knee joint forces using wearable sensors during sixteen different sports movements, demonstrating correlation coefficients that ranged from 0.64 to 0.90 for anterior-posterior knee joint forces. The results showed good agreement between the ANN-predicted knee joint forces and actual biomechanical measurements indicating the potential of wearable technology combined with ANN in sports injury prevention and management. However, the computational resources are much greater for ANN compared to XGBoost. ANNs generally require more memory and processing power due to their complex structures with millions of parameters and computationally intensive training involving deep layers and matrix multiplications. XGBoost is more memory-efficient and computationally faster, utilizing simpler structures, gradient boosting techniques, and CPU-based parallelization, making it less demanding on resources. This comparison underscores the trade-off between the predictive accuracy of ANN models and the more resource-efficient alternatives like XGBoost, highlighting a critical consideration in the selection of technology and data processing for desired biomechanical analysis.

This research has several limitations, including a small, homogeneous sample of 23 healthy, recreationally active females, limiting its broader applicability. The reliance on IMUs for data collection could introduce variability, affecting precision. Additionally, focusing solely on drop jump movements may not accurately represent the ACL's experiences across various sports and activities, potentially limiting the model's real-world predictive accuracy.

Future work should aim to validate the refined model across a broader population, encompassing a range of athletic abilities and movement patterns. Additionally, exploring the integration of machine learning models with real-time feedback systems could pave the way for immediate and actionable insights during athletic training and rehabilitation practices. Lastly, this research supports the continued exploration into the capacities of wearable technology in sports science, particularly as a tool for injury prevention and performance optimization.

CONCLUSION: This research successfully demonstrates that machine learning models, specifically XGBoost utilizing IMU data, can effectively predict ASF, offering a non-invasive tool for injury prevention. The streamlined model with reduced features and signal filtering shows high predictive accuracy, making it a viable option for quick feedback in athletic scenarios. This advancement paves the way for broader application and could significantly impact the fields of sports science and rehabilitative medicine.

REFERENCES

Beaulieu, M. L., Ashton-Miller, J. A., & Wojtys, E. M. (2023). Loading mechanisms of the anterior cruciate ligament. *Sports Biomechanics, 22*(1), 1-29. https://doi.org/10.1080/14763141.2021.1916578

Chappell, J. D., Yu, B., Kirkendall, D. T., & Garrett, W. E. (2002). A comparison of knee kinetics between male and female recreational athletes in stop-jump tasks. *Am J Sports Med, 30*(2), 261-267.

Fong, D., & Chan, Y.-Y. (2010). The Use of Wearable Inertial Motion Sensors in Human Lower Limb Biomechanics Studies: A Systematic Review. *Sensors, 10*(12), 11556-11565. https://doi.org/10.3390/s101211556

Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2019). Estimation of Knee Joint Forces in Sport Movements Using Wearable Sensors and Machine Learning. *Sensors, 19*(17), 3690. https://doi.org/10.3390/s19173690

Weinhandl, J. T., Joshi, M., & O'Connor, K. M. (2010). Gender Comparisons between Unilateral and Bilateral Landings. *Journal of applied biomechanics, 26*(4), 444-453. https://doi.org/10.1123/jab.26.4.444