ACTA GYMNICA, 2025, Volume 55, Article e2025.007
https://doi.org/10.5507/ag.2025.007

OPEN ACCESS

ORIGINAL RESEARCH

Can we predict archery performance using shooting heart rate and other biophysical parameters? A machine learning approach

Chandra Sekara Guru[1]✉, Uma Mahajan[2], Anup Krishnan[3], Karuna Datta[1], and Deep Sharma[1]

[1]Department of Sports Medicine, Armed Forces Medical College, Pune, India; [2]Army Sports Institute, Pune, India; and [3]School of Sports, Exercise and Nutrition Sciences, DY Patil University, Navi Mumbai, India

Abstract

Background: Heart rate (HR) during different shooting phases, along with other biophysical parameters, is used for performance analysis. Given the stochastic and complex nature of these parameters, machine learning (ML) models have recently been used in predicting performance. **Objective:** We aimed to develop an ML model to predict the archery performance score using key performance parameters, including the HR values during key phases of shooting, and determine the predictive importance of these parameters to provide personalised training to an archer. **Methods:** A total of 32 archers (15 elite and 17 non-elite) shot two sessions of 30 arrows each indoors, wearing an HR chest monitor and were videographed. When each arrow was shot, 11 HR values were identified at different shooting phases. Biophysical parameters consisting of 35 linear variables and second-degree polynomial HR values were used to build ML models in Python. The total scores of Sessions 1 and 2 were used to train and test, respectively. Model performance was evaluated using root mean squared error and highest accuracy. Shapley's additive explanation method was used to show the predictive importance of each variable. **Results:** The Cat Boost ML model with the lowest error and better accuracy was used to predict the archery score. Sports age, resting systolic blood pressure, previous competition score, right-hand grip strength, age, HR rate before 2 s of arrow release, waist-to-hip ratio, concentration disruption, trait anxiety and HR after 5 s of release are the top parameters that predicted the score. **Conclusions:** This study highlights that the ML approach can be a useful tool to predict top parameters to optimise archery performance. The ML model has been trained and tested for indoor archery settings in a controlled environment, which may be a limitation in generalising this ML model for prediction during outdoor competitive settings.

Keywords: machine learning, athletic performance, sports performance, heart rate control, sports training

Inroduction

Sports performance in archery is the ability of the archer to shoot the maximum score during the event repeatedly (Kolayiş & Mimaroğlu, 2008). The complex interaction of the physiological, psychological and physical factors results in this coordinated motor performance. Adaptations in these systems resulting from periodised scientific sports training can optimise the shooting performance in archery. The evaluation of the athletes periodically during the training phases and analysis of these biophysical performance parameters by the coach and performance staff are key to gaining insight to undertake corrective action. These biophysical parameters have been used to predict performance in archery.

One such performance parameter is the shooting heart rate (HR), an outcome effect of the adaptive responses of the physiological and psychological systems. Most studies observed a correlation between the HR values before shooting and the archery scores (Aggarwala & Dhingra, 2017; Clemente et al., 2011). Archery shooting involves different key phases, that is, aiming, release, and follow-through phases. HR values and other biophysiological parameters recorded during these phases significantly correlated with archery scores (Keast & Elliott, 1990; Mohamed et al., 2014; Soylu et al., 2006). Even in the Tokyo 2020 Olympics, the real-time HR of archers was recorded and displayed on the screen during the archery competitions. Lu and Zhong (2023) examined the relationship between HR values and archery scores during the Tokyo Olympics. They found that higher HR values are associated with lower performance scores. When controlled for individual and match level, this correlation was robust. Hence, in our study, we intended to capture the HR values during these key phases for deeper insights. Despite this correlation, shooting HR values alone without the other performance factors may not aid in performance prediction and training optimisation (Aggarwala & Dhingra, 2017; Guru et al., 2020; Kolayiş & Mimaroğlu, 2008; Lu & Zhong, 2023; Robazza et al., 1999).

Every athlete is unique, given the differences in genetic composition, gender, anthropometric characteristics and psychological traits (Harvey et al., 2020; Sellami et al., 2021). Specific determinants and motor skills required

✉ **Corresponding author:** Chandra Sekara Guru, e-mail drscsguru@gmail.com, ORCID® record https://orcid.org/0000-0001-6032-4689

Article history: Received July 17 2024, Accepted July 31 2025, Published August 19 2025

according to the sport discipline further increase the complexity. Moreover, the parameters that are periodically evaluated are closely interrelated and stochastic. So, conventional statistical models using linear or logistic regression are difficult to apply in such conditions with high multicollinearity. Hence, research studies incorporating machine learning (ML) regression models have been postulated instead of routine statistical regression methods to predict the archers' performance (Mahajan et al., 2021; Muazu Musa et al., 2019a, 2019b; Taha, Musa, Abdul Majeed, Abdullah, & Hassan, 2018).

In the last decade, anthropometric characteristics, fitness parameters, physiological resting variables and biomechanical variables have been used to develop ML prediction models to either identify or classify the performance potential of the archers (Mahajan et al., 2021; Muazu Musa et al., 2019a, 2019b; Taha, Musa, Abdul Majeed, Abdullah, Ab. Nasir, et al., 2018; Taha, Musa, Abdul Majeed, Abdullah, & Hassan, 2018). Based on the fitness and motor ability performance variables, researchers have employed ML models not only in the identification of potentially talented archers but also in the classification of high-performance archers (Muazu Musa et al., 2019a; Taha, Musa, Abdul Majeed, Abdullah, Ab. Nasir, et al., 2018; Taha, Musa, Abdul Majeed, Abdullah, & Hassan, 2018). Mahajan et al. (2021) used physical fitness parameters and physiological parameters to compare ML models and conventional statistical methods for selecting important features that predict archery performance. However, none of the existing predictive ML models have incorporated shooting HR values for archery performance prediction, which has been found to have a high real-time correlation. There is a lack of a performance predictive ML model that incorporates both key performance parameters and the HR values captured during shooting. These limitations and gaps in the existing ML models give rise to research questions, namely, Does an ML model incorporating both the shooting HR values and the key performance variables predict archery performance score? Can such a model identify the predictive importance of each of these variables, specific to a particular archer? Such an understanding of the individual parameter's predictive importance is key to optimising performance in a specific archer and can aid the coach in modulating the training methods to provide individualised performance enhancement.

Hence, this study aimed to develop an ML model to predict the archery performance score using key performance parameters, including the HR values during key phases of shooting, and determine the predictive importance of these parameters to provide personalised training to an archer.

Methods

An observational study was designed among the male archers of a residential sports institute in India after obtaining the necessary ethical clearance from the Institutional Ethical Committee of the Armed Forces Medical College, Pune (IEC/NOV/2016). A total of 32 Indian archers (15 elite and 17 non-elite) with a mean age of 21.8 ± 2.0 years volunteered to take part in this study after providing informed consent. The group consisted of 21 recurve and 11 compound archers. All participants were in active sports training and participated in the absence of any audience or coaching personnel. The archers underwent common periodised training with a weekly micro-cycle consisting of general and sports-specific physical fitness training, archery shooting skills training, mental training sessions and active recovery sessions.

On the day of the study protocol, body weight and height were measured using a scale and a stadiometer (DHRWM, Cardinal Scales Manufacturing, Webb City, MO, USA) to the nearest 0.1 kg and 0.1 cm, respectively. Waist and hip girth were measured in centimetres using a non-extensible anthropometric measuring tape by the same investigator (Marfell-Jones et al., 2006). Waist-to-hip ratio was calculated from these measured values. After recording resting HR and blood pressure in the right arm in the sitting posture, body fat percentage was assessed by bio impedance analysis (Bodystat Quadscan 4000, Sulby, Isle of Man, British Isles). The average of the three readings of the hand grip strength of both hands measured using a digital dynamometer (SH5001, Saehan Co., Gyeonggi-do, South Korea) was used. Archers also filled out a questionnaire on sports participation, previous night sleep quality and the Sports Anxiety Scale (Smith et al., 1990). These biophysical parameters were used in the ML model to predict archery performance.

The study consisted of 10 ends each in 2 sessions, and 30 arrows in each session were shot from a distance of 18 m in an indoor archery hall as per the World Archery Association standard rules (Guru et al., 2020). The demographic and sports profiles of the participants are presented in Table 1. The archers were grouped as elite compound (21.9%), elite recurve (25%), non-elite compound (12.5%) and non-elite recurve (40.6%). About 75% of these archers were training in the preparatory phase, and 25% were in the pre-competitive phase during the conduct of the study. The principal investigator recorded the entire shooting process and the corresponding scores made by each arrow. The environment set-up, timing of the study and pre-participation requirements were the same for all the participants. The participants were advised to sleep at least 7–8 hours the previous night, avoid vigorous physical activities the previous day, avoid caffeinated drinks and heavy meals before the study protocol, and adhere to their daily routine. HR before performing the warm-up activity and during the shooting activity was recorded for all the participants. HR was measured using an HR monitor with a chest belt (H7, Polar Electro Oy, Kemple, Finland). The wristwatch displaying the HR was fixed to a tripod and filmed simultaneously while shooting during the observed period. The video recorded was set at 1920 × 1080 at 60 frames per second (HDR-PJ670, Sony, Tokyo, Japan).

Slow-motion analysis of the recordings was done to identify the time of release of the arrows and frames, 5 s before and 5 s after the arrow release. All the frames had the shooting HR value displayed on the watch. Thus, a total of 11 shooting HR values covering the aiming, release and

Table 1 Profile of Indian archers (N = 32)

Variable	Value
Anthropometric, *M* (*SD*)	
Age (years)	21.8 (5.63)
Sports age (years)	7.6 (5.47)
Training age (years)	4.1 (3.63)
Height (cm)	168.7 (6.35)
Weight (kg)	66.7 (9.80)
Waist-to-hip ratio	0.9 (0.04)
Body fat (%)	10.9 (2.88)
Physiological, *M* (*SD*)	
Right hand grip strength (kg)	41.6 (7.08)
Left hand grip strength (kg)	40.8 (7.22)
Resting heart rate (beats per minute)	58.8 (7.62)
Resting blood pressure – systolic (mmHg)	114.5 (9.17)
Resting blood pressure – diastolic (mmHg)	64.4 (7.33)
Heart rate before warm-up (beats per minute)	88.9 (10.64)
Psychological, *M* (*SD*)	
Somatic trait anxiety (9–36)	13.9 (2.68)
Cognitive trait anxiety (7–28)	13.1 (3.93)
Concentration disruption trait anxiety (5–20)	8.2 (2.12)
Sleep quality, *n* (%)	
Good	25 (78.1%)
Average	7 (21.9%)
Highest performance in previous competition, *M* (*SD*)	
Archery score (out of 720)	660.4 (29.68)
Type of previous highest performing competition, *n* (%)	
International	7 (21.9%)
National	11 (34.4%)
Below national	14 (43.8%)
Outcome (archery score out of 300), *M* (*SD*)	
Session 1	279.4 (7.43)
Session 2	279.5 (9.76)

follow-through phases were captured for each arrow. The detailed materials and methods have already been presented in the pilot study (Guru et al., 2020). With a total of 32 archers shooting 60 arrows each and 11 shooting HR values, the total HR samples used in this study to develop and test the ML performance predictive model were 21,120. The mean HR values were plotted against the 11 frames depicting the three archery phases: aiming (–5 s to –1 s), release (0 s), and follow-through (+1 s to +5 s) to identify the pattern.

Descriptive statistics, namely frequencies, percentages, mean and standard deviation, along with 95% confidence intervals, were used to summarise categorical and numerical data, respectively. A correlation heat map was used to depict the correlation between numeric variables. The independent predictor variables are tabulated in Table A1. As we mentioned above, thirty-two archers shot 30 arrows in two sessions, resulting in a total of 1,920 arrows shot. This data was used as the training and testing dataset, respectively. The maximum score for each shot was 10, so the maximal archery score for the session was 300.

Due to the curvilinear nature of the shooting HR values observed when plotted against time (Figure 1), we used two approaches to develop a total of 12 ML regression models. A total of 35 variables, comprising profile of archers, anthropometric, biophysical, psychological, and 11 HR values (–5 s, –4 s, …, 0 s, ..., +4 s, +5 s arrow release) were included as the independent variables. In the first approach, we assumed all 11 HR values as linear variables and developed 6 ML models. In the second approach, we also considered the second-degree polynomials of these 11 HR values to develop another 6 ML models. The details of variables are provided in Table A1. Session 1 data (n = 960) was used to train the ML models. These ML Models were fine-tuned using Grid Search CV cross-validation. The predictive performance of each ML model was tested using the Session 2 score dataset (n = 960). root mean squared error (RMSE), R^2 score, mean absolute percentage error (MAPE), and accuracy ([1 – MAPE] * 100) were derived for evaluating the model performance. The ML model with the lowest RMSE and highest accuracy was considered the final model.

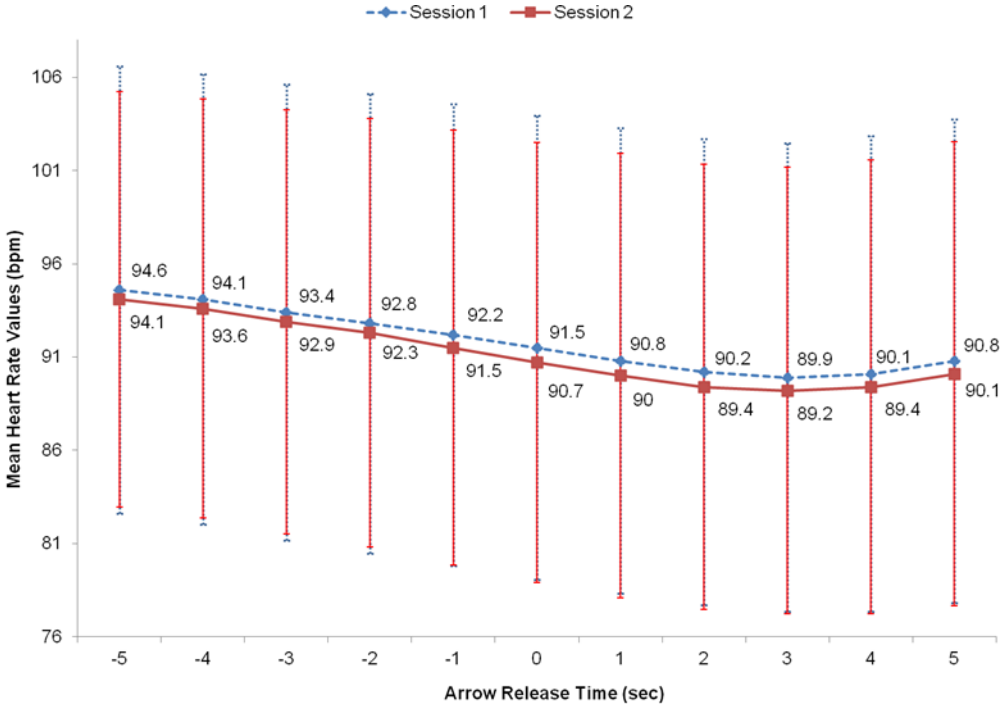
Variable importance is useful to identify and select important variables to develop the ML model. It measures the overall contribution of each feature to the model's performance across all predictions. However, it doesn't explain why a specific prediction is made by the model. SHapley Additive exPlanations (SHAP) method is useful to understand the predictions made by any ML model. It is used for all types of models from linear regression to deep learning, hence known as "model-agnostic". It gives a more detailed, game-theory-based explanation of each variable contributing to overall (global) and individual data point-level (local) predictions (Lundberg & Lee, 2017). It also explains the quantity and direction of each variable/feature's contribution to increasing or decreasing the predicted outcome. Hence, SHAP analysis was carried out to identify drivers of archery score and quantify their contributions to predicting archery score at an individual-archer level. All the analysis and ML modelling were done in Python v3.11.4.

Results

The mean shooting HR values identified during the aiming, release and follow-through phases showed a specific deceleration pattern in both sessions (Figure 1). Maximum deceleration of the mean shooting HR values was observed during +2 s and +3 s after the release of the arrows in the follow-through phase, and then gradually started to increase further. Even though the mean HR values were lower in Session 2 compared to Session 1, the difference in mean archery scores between the sessions was not statistically significant.

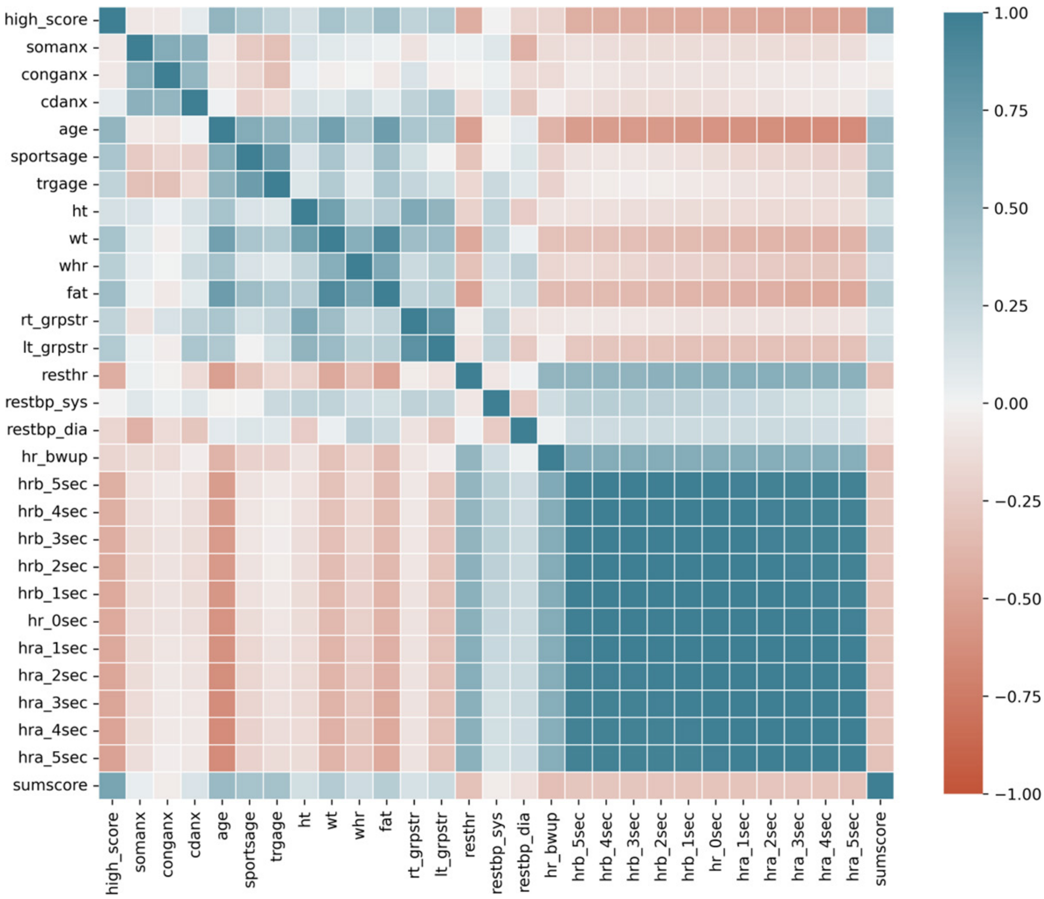
The correlation heatmap of predictor variables and the archery score is shown in Figure 2. The blue and red colour boxes represent the positive and negative correlation, respectively. The lighter to darker shade represents the strength of the correlation from low to high. The highest archery score in the previous competition, age of the archer, training age at the institute, and sports age (experience) were positively correlated with the Session 1 archery score. The shooting HR values of all three key phases negatively correlated with the archery score.

Figure 1 Heart rate values in archers



Note. Error bars represent standard deviations.

Figure 2 Correlation heatmap of predictors variables and archery scores



Note. Variables are described in Table A1.

Table 2 shows the ML models developed using both the approaches and their respective RMSE, R^2 score, MAPE and accuracy along with the best model parameters.

The linear Categorical (Cat) Boost ML model was chosen as the final model with the lowest RMSE of 6.262 among the 12 developed models after testing with the Session 2 dataset. The Cat Boost linear ML model developed using the Session 1 dataset was used to predict the archery score and was tested with the actual Session 2 archery score of all the archers individually (Figure 3). The average predicted score using the Cat Boost linear ML model was 279.4. Figure 3 depicts the difference in the actual and predicted archery scores as well as the deviation away from the average predicted score between the category groups of the archers, especially between elite and non-elite archers.

Predictor importance of the biophysiological variables

Figure 4 shows the summary plot of the SHAP values derived using the Cat Boost ML model with linear shooting HR values. The SHAP values depict the overall contribution of the independent variables in predicting the archery performance score in the Session 2 dataset (Figure 4). The predictor/independent variable names are shown on the Y-axis, and SHAP values of each variable are shown on the X-axis in the plot. The variables are displayed from top to bottom in order of their importance, from most important to least important in predicting the archery score, respectively. The SHAP values from low to high are shown using blue and red colours, respectively. The lighter to darker shade represents the magnitude of the SHAP values from low to high. Each point on each predictor variable represents the SHAP value of each archer.

Table 2 Machine learning model performance in testing dataset

Model	RMSE	R^2 score	MAPE	Accuracy (%)	Parameter
Without polynomial features (35 variables)					
Linear regression	7.489	0.393	0.022	97.792	
ElasticNet	7.606	0.374	0.022	97.802	alpha: 25, l1_ratio: 0.5
Random forest	6.919	0.482	0.019	98.065	max_features: 5
XG Boost	6.728	0.510	0.019	98.054	learning_rate: 0.15, max_depth: 2, n_estimators: 40
Cat Boost	6.262	0.575	0.018	98.180	learning_rate: 0.5, max_depth: 4, n_estimators: 25
Light GBM	9.612	0.000	0.027	97.271	learning_rate: 0.1, max_depth: 2, n_estimators: 20
With polynomial features (46 variables)					
Polynomial regression	8.325	0.250	0.023	97.711	
ElasticNet	7.394	0.408	0.022	97.849	alpha: 15, l1_ratio: 0.5
Random forest	6.852	0.492	0.019	98.064	max_features: 6
XG Boost	6.728	0.510	0.019	98.054	learning_rate: 0.15, max_depth: 2, n_estimators: 40
Cat Boost	7.334	0.418	0.021	97.899	learning_rate: 0.15, max_depth: 2, n_estimators: 20
Light GBM	9.612	0.000	0.027	97.271	learning_rate: 0.1, max_depth: 2, n_estimators: 20

Note. RMSE = root mean squared error; MAPE = mean absolute percentage error; accuracy = (1 – MAPE) * 100; GBM = Gradient Boosting Machine.

Figure 3 Session 2 Actual and predicted archery scores (Cat Boost model – 35 variables)

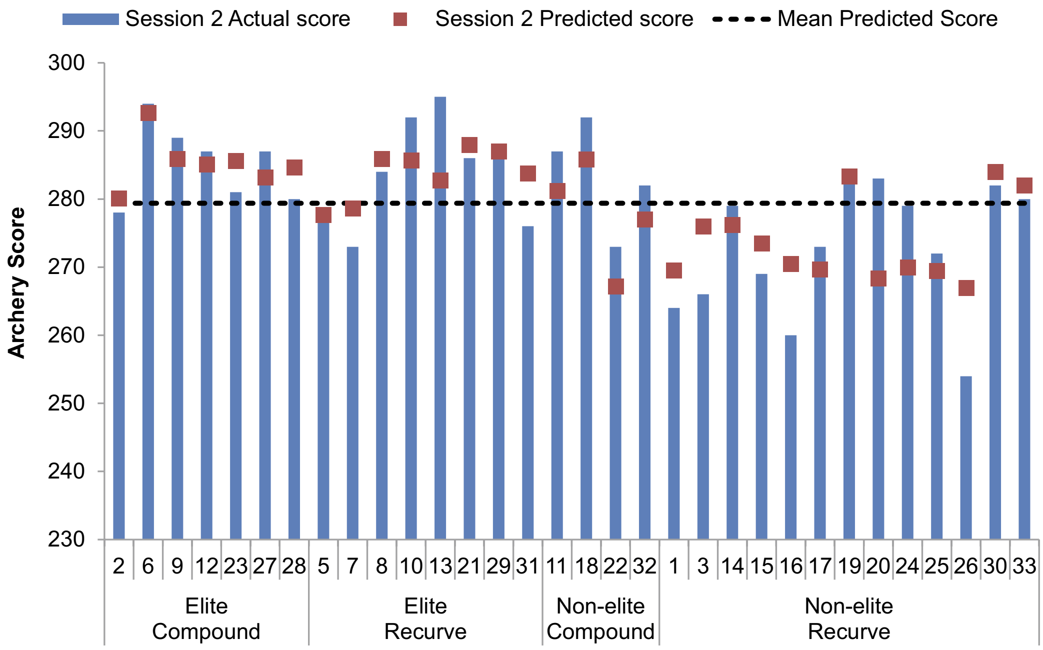
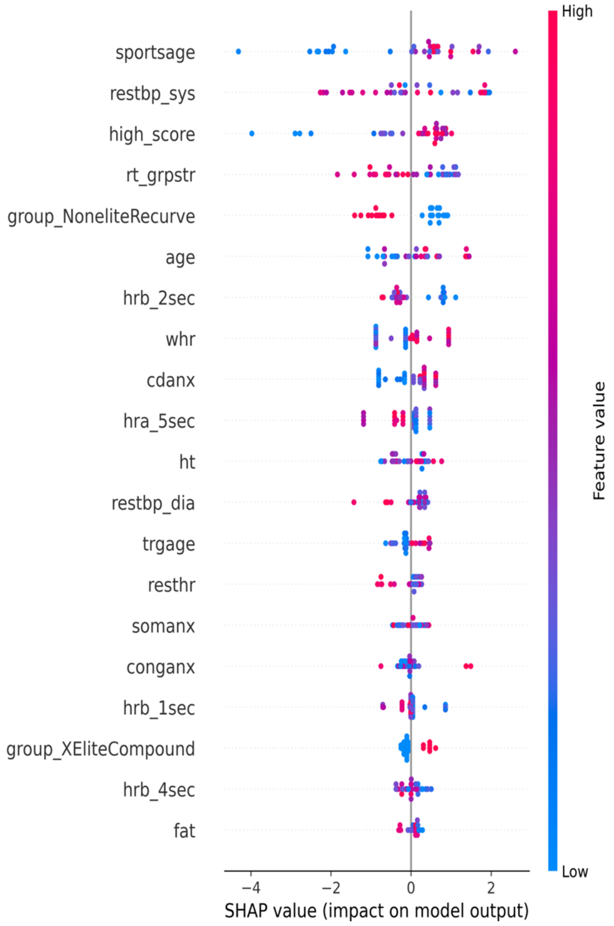


Figure 4 Predictor importance of Cat Boost Model using SHAP analysis



Note. Variables are described in Table A1.

Sports age, resting systolic blood pressure, high archery scores in previous competitions, HR at 2 s before and 5 s after arrow release, concentration disruption trait anxiety (5–20), and the non-elite recurve archer group were the most important predictors. The higher values of sports age, training age, archer's biological age, previous competitions archery score and concentration disruption trait anxiety (5–20) were contributing to increasing the predicted archery score. Whereas, the lower values of resting systolic/diastolic blood pressure, right-hand grip strength, HR 2 s before and 5 s after arrow release, and resting heart rate were contributing to predicting a higher archery score. Predicted archery score was lower in the non-elite recurve archers group and higher among the elite compound archers group.

A comparison has been drawn to highlight how the contribution of the performance variables varies between two archers despite belonging to the same archery bow type (recurve vs compound) and competitive level (elite vs non-elite). Figure 5 shows the difference in the predictor variables contribution using SHAP analysis waterfall plots of the Cat Boost ML model between archers from the same categories, that is, elite compound (archer ID 6 vs. archer ID 23) and non-elite recurve (archer ID 20 vs. 19). These plots explain the contribution of each variable in increasing/decreasing the individual's predicted score from base value, that is, the mean archery score. The values inside the red arrows of the corresponding variable push the model towards predicting a higher predicted score. In contrast, those in blue arrows push the model towards the low predicted score. The range of the predicted archery scores is displayed on the X-axis, and the mean archery score E[f(x)] = 279.382 is at the bottom of the X-axis, and the predicted archery score is at the top of the X-axis. The contribution of variables in predicting archery score for individual archers is calculated using a function:

Predicted archery score = Mean archery score + sum(SHAP values of all variables).

Discussion

We aimed to measure the HR during the key phases of shooting using basic field-side aids to highlight that coaches and scientists can easily incorporate this method during routine periodic evaluations (Guru et al., 2020). Our attempt to use multiple variables, including shooting HR to develop a performance prediction ML model for the athlete's personalised training, is a novel approach.

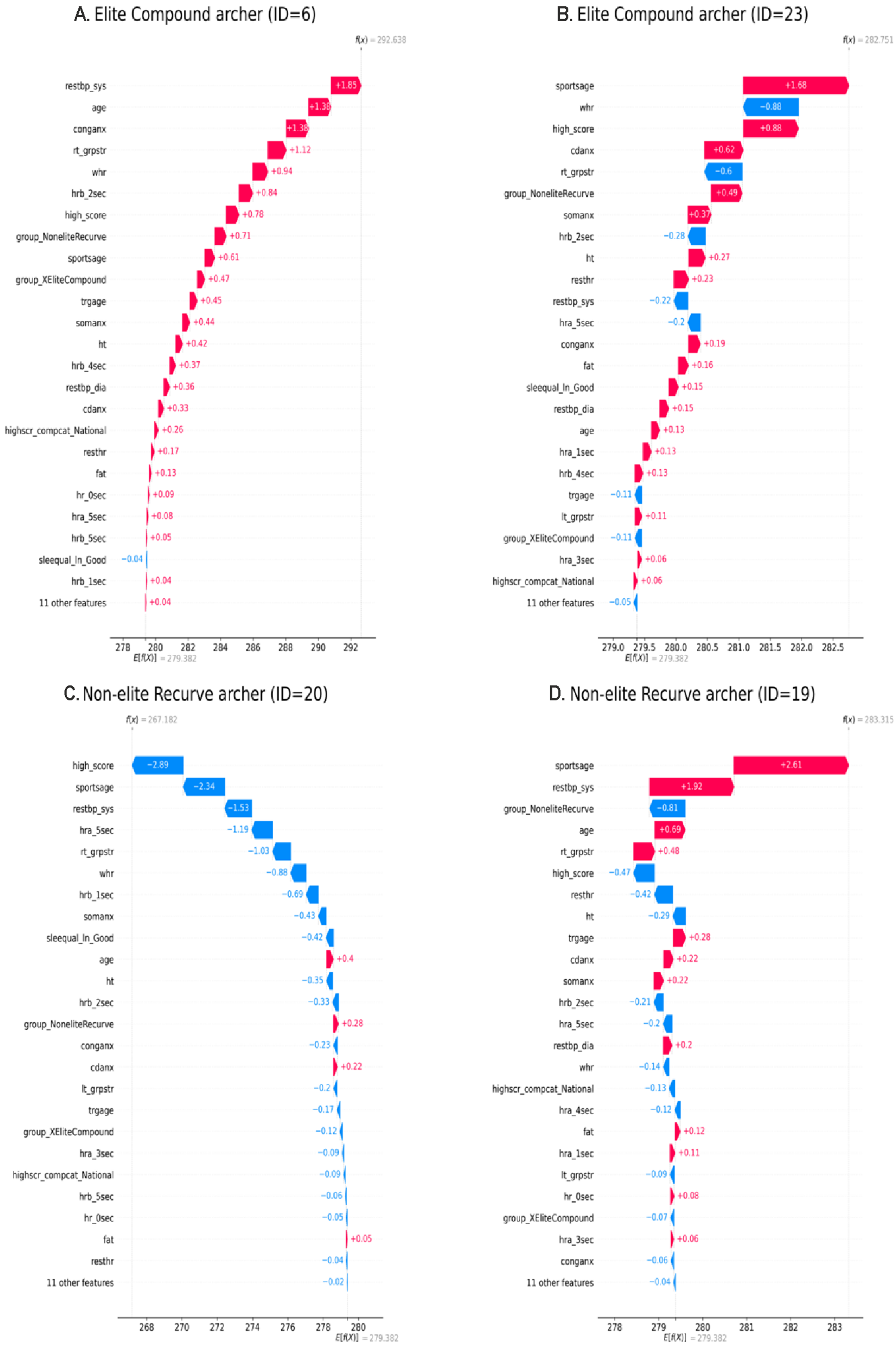
The performance variables grouped under physical, mechanical, physiological and psychological components are varied. These factors vary for different events, and their component contribution also may differ between individual athletes (Kim et al., 2015). The individual archer may develop unique strategies and adaptations to training in these performance components (Vendrame et al., 2024). Hence, we included four categories of archers based on competitive level (elite vs non-elite), and archery event (recurve vs compound) in our study to gain deeper insight. A total of 35 variables under anthropometric, sports-related, physiological, psychological, sleep and shooting HR values during the key shooting phases were included (Table A1) based on previous research studies and correlation with archery performance scores (Aggarwala & Dhingra, 2017; Clemente et al., 2011; Datta et al., 2020; Guru et al., 2019, 2020; Kim et al., 2015; Mahajan et al., 2021).

Pattern of shooting HR values

A unique feature selection variable to predict performance included in this study was shooting HR values during the three main shooting phases. Clemente et al. (2011) and Aggarwala and Dhingra (2017) found an inverse correlation between pre-shooting HR values and archery scores. Robazza et al. (1998) identified a specific deceleration pattern of HR during the transition from draw-aiming to release phase, which correlated with the archery scores and level of archers. Lu and Zhong (2023) used the real-time shooting heart rate and archery score telecasted during the 2020 Tokyo Olympics, also found an inverse relation. This inverse relation of higher HR and lower scores was robust when controlled for fixed effects at archer and match. We found that the shooting HR values showed a specific deceleration pattern that extended beyond the release phase into the follow-through phase in our study (Figure 1).

The deceleration pattern in shooting HR values has been attributed to the outcome effect of coordinated

Figure 5 Individual archer-level importance of predictors



Note. Variables are described in Table A1.

motor, breathing, attention and visual responses during the aiming-shooting phases due to cardiac-somatic coupling, intake-rejection hypothesis or the quiet eye phenomenon in aiming sports (Gonzalez et al., 2017; Neumann & Thomas, 2009; Robazza et al., 1998). Güler et al. (2024), also in their experimental study, found a similar HR deceleration pattern before shooting with prolonged quiet eye duration and higher scores among intermediate air pistol shooters as compared to their novices. The HR deceleration has been correlated with other physiological parameters such as the electroencephalogram and breathing pattern in aiming sports (Neumann & Thomas, 2009; Wang et al., 2023). Being an easy-to-record outcome parameter of multiple physiological phenomena as well as a field-side tool, HR deceleration can thus provide better clues to coaches for planning sports training. The deceleration pattern when plotted against time was curvilinear (Figure 1). Hence, the second approach includes second-degree polynomials of mean shooting HR values to develop the other 6 ML models.

ML model approach for performance prediction

Traditional statistical models are based on stringent assumptions like a linear relationship, normally distributed errors, and homoscedasticity. Performance components are multiple, multicollinear, and interdependent (Mahajan et al., 2021). This is highlighted in the correlation heatmap plot (Figure 2). In archery, parameters like archers' age, sports age or training years in archery and the previous competition high score are positively correlated to performance at varying degrees represented in the heat map. This indirectly reveals the skill acquisition and confidence the archer develops (Kim et al., 2015). Parameters like resting heart rate, resting blood pressure and heart rate before the warm-up are negatively correlated with performance due to the physical and cognitive effects of training on the parasympathetic system in archers (Aggarwala & Dhingra, 2017; Lee, 2009; Muazu Musa et al., 2019a).

Traditional statistical regression methods filter these crucial, complex variables during feature selection analysis and hence may have limitations in performance prediction. In recent studies, ML models have been used to tackle this complexity to predict sports performance and identify potentially talented athletes (Etxegarai et al., 2018; Hoog Antink et al., 2021; Mahajan et al., 2019; Muazu Musa et al., 2019a; Taha, Musa, Abdul Majeed, Abdullah, & Hassan, 2018). Recently, an ML model has been developed with 90% reliability to classify archers based on archery performance using a single physiological parameter, like quiet eye duration during aiming (Söğüt et al., 2025).

In our study, we used two different approaches to develop 12 different ML models. RMSE ranged between 6.262 and 9.612 for all models. Despite including second-degree polynomials, the Cat Boost ML model with linear HR values itself predicted the archery score better with the lowest RMSE of 6.262 (Table 2). It was used to predict the Session 2 score and tested against the actual Session 2 scores (Figure 3), which is again an important strength of this study, given the large data sets for both training and testing.

Cat Boost linear ML model predicted a baseline archery score of 279.38 out of 300 in Session 2 across all groups of archers. It outperformed (RMSE 6.262; accuracy 98.18%) compared to other ML models, especially for categorical variables group (elite/non-elite, compound/recurve), training phase (preparatory/pre-competition), sleep quality last night (average/good), and type of previous competitions – highest score (below national/national/international). Other tree-based models, namely random forest (RMSE 6.919), eXtreme Gradient Boosting (XGBoost; RMSE 6.728) and light Gradient Boosting Machine (GBM; RMSE 9.612) displayed poor performance while dealing with categorical variables. An interesting aspect is the difference in deviation of individual predicted and actual scores from the baseline predicted score noted between the archer groups. Among most of the non-elite recurve archers, both actual and predicted scores were below the baseline predicted score, which might have been due to the difference in the learning curve and comparatively tougher equipment handling technique involved in recurve archery (USA Archery, 2012). This aspect may help the coach plan the training program accordingly.

Importance of the performance parameter in predicting the score

The baseline predicted score alone may only give a general idea to the coach and the athlete for further performance enhancement. The relative contribution of various individual variables to the predicted score can aid in personalised training. Hence, SHAP analysis on the Cat Boost model was done to identify the contribution of each independent variable in predicting the performance (Figure 4). Sports age, resting systolic blood pressure, highest archery score by the archer in previous competitions, right hand grip strength, non-elite recurve group, age, HR value 2 s before arrow release, waist-to-hip ratio, concentration disruption trait anxiety, and HR value 5 s after arrow release were the top 10 variables which played a vital role in the prediction of the archery score.

An experience of five years in archery is emphasised for acquiring the mental skills to perform better by Korean archers (Kim et al., 2015). Our ML model also highlights the importance of sports age in predicting archery performance. Similar to our findings, Muazu Musa et al. (2019b) and Taha et al. (2018) have also highlighted the lower resting blood pressure (BP), resting HR, and higher handgrip strength as performance indicators of high-performing potential archers. Mahajan et al. (2021) found that maximum handgrip strength, higher fat mass, and age are important features in their ML feature selection analysis for archery performance. Right-hand grip strength is one of the important variables to predict archery performance in other experimental studies conducted in archers, mainly due to the load required each time to pull the bow back during shooting (Handayani et al., 2024; Rusdiawan et al., 2024).

Our study, in addition to these variables, has identified concentration disruption trait anxiety, a component of the Sports Anxiety Scale. The concentration disruption trait score ranges between 5 and 20, and a higher score is negatively correlated with performance (Smith et al., 1990). In our study, the mean score was only 8.2 with a 95% confidence interval of 7.45 to 8.98 (Table 1). Shaari et al. (2023), in their study among 44 national-level archers, also found a similar concentration disruption mean score of 10.33.

SHAP analysis further amplified the importance of the shooting HR values during the follow-through phase at 2 s and 5 s after the release of the arrows in predicting the archery score. This is an important input for coaches to focus on during this period in the training sessions. Also, sports-related variables like sports age (experience), previous high scores in competitions, and archer belonging to a non-elite recurve group are key factors that predict performance.

Practical application for personalised training

The contribution of the predictor variables for each archer in predicting the archer's performance varies. Information on the contribution of these variables to a particular archer will provide better insight for the coaches to plan personalised training and enhance their performance. A comparison of how the variables contributed to the actual score from the baseline predicted score between two archers of the same group, namely Elite compound (Figure 5A and 5B) and Non-elite recurve (Figure 5C and 5D), is depicted.

In this first instance, we compared elite compound archer ID 6 vs elite compound archer ID 23, a scenario where the individual predicted scores of both archers were better than the average ML predicted score of 279 for all archers, that is, 292 and 282, respectively. When we further analyse the specific contributions by the variables towards this better predicted score using the ML model, the resting systolic BP contributed approximately 2 points (+1.85) in archer ID 6 (Figure 5A), whereas the sports age (experience) contributed approximately 2 points (+1.68) in archer ID 23 (Figure 5B). Thus showing the different performance variables contributing to the performance prediction among archers of the same group. Further, performance variables like age, cognitive trait anxiety, right-hand grip strength, waist-to-hip ratio, and HR 2 s before arrow release improved the predicted score by 1 point each in archer ID 6 (Figure 5A). Whereas in archer ID 23, SHAP analysis shows that waist-to-hip ratio (–0.88), right grip strength (–0.6) and 16 other features displayed a negative impact on the individual predicted score (Figure 5B). This individual archer-based insight, especially with the weightage contribution of the performance variables, shall aid the coach in tweaking the training program. In this particular archer ID 23, the coach may pay attention to the body composition, upper body exercises to improve grip strength and breathing or mental focus exercises during the aiming and follow-through phases to address the highlighted factors.

In the second instance, we compared archer ID 20 vs archer ID 19, who were both non-elite recurve archers in a scenario where the individual predicted score of one archer was lower (ID 20) and the other archer was higher (ID 19) than the average predicted score of 279. In archer ID 20, the previous competition performance, sports age, resting systolic BP, HR value 5 s before arrow release, and right-hand grip strength played a major role in decreasing the individual predicted score by approximately 9 points from the average predicted score of 279 (Figure 5C). Whereas in the case of archer ID 19, the individual's predicted score further improved to 283 from the average predicted score of 279. Here, sports age and resting systolic BP were vital in improving the predicted archery score by 3–4 points (Figure 5D). Thus, within the same non-elite recurve archers group, it can be seen that the contributions of the predictor features differ, which may be a clue to the coach for planning personalised training. Given the individual genetic and phenotypic differences among the archers, such individual-specific insights about performance factors early in the sports career also help the coach to give personalised attention to a particular factor during sports training.

Limitations

In our model, we have not included other archery-specific fitness variables. The inclusion of physical fitness variables in the model in future may increase the accuracy of prediction. The ML model has been trained and tested for indoor archery settings in a controlled environment, which may be a limitation in generalising this ML model for prediction during outdoor competitive settings. Future studies in competition outdoor settings may pave the way for better, accurate and reliable ML models for archery performance prediction.

Conclusions

The Cat Boost ML model with the lowest error and better accuracy was used to predict the archery score using the shooting HR values and other biophysical parameters. Sports age, resting systolic blood pressure, previous competition score, right-hand grip strength, age, heart rate before 2 s of arrow release, waist-to-hip ratio, concentration disruption trait anxiety and HR after 5 s of release are the top parameters that predicted the score. This study highlights that the ML approach can be a useful tool to predict top parameters to optimise archery performance.

Acknowledgments

The authors acknowledge the support and active involvement of the coaches, support staff and administrators of the Army Sports Institute, Pune, India during the study.

Conflict of interest

The authors report no conflict of interest.

References

Aggarwala, J., & Dhingra, M. (2017). Effects of autonomic control on performance of archers: A comparative study on novice and experienced archers. *International Journal of Biomedical Research, 8*(4), 182–186.

Clemente, F., Couceiro, M., Rocha, R., & Mendes, R. (2011). Study of the heart rate and accuracy performance of archers. *Journal of Physical Education and Sport, 11*(4), 434–437. https://efsupit.ro/images/stories/imgs/JPES/2011/12/10Art_66.pdf

Datta, K., Kumar, A., & Sekar, C. (2020). Enhancement of performance in an elite archer after non-pharmacological intervention to improve sleep. *Medical Journal Armed Forces India, 76*(3), 338–341. https://doi.org/10.1016/j.mjafi.2018.06.010

Etxegarai, U., Portillo, E., Irazusta, J., Arriandiaga, A., & Cabanes, I. (2018). Estimation of lactate threshold with machine learning techniques in recreational runners. *Applied Soft Computing, 63,* 181–196. https://doi.org/10.1016/j.asoc.2017.11.036

Gonzalez, C. C., Causer, J., Grey, M. J., Humphreys, G. W., Miall, R. C., & Williams, A. M. (2017). Exploring the quiet eye in archery using field- and laboratory-based tasks. *Experimental Brain Research, 235*(9), 2843–2855. https://doi.org/10.1007/s00221-017-4988-2

Güler, G., Şimşek, D., Tuncer, S., & Mon-López, D. (2024). Evaluation of physiological parameters and gaze behaviour of air pistol shooters related to performance outcomes. *International Journal of Performance Analysis in Sport, 24*(3), 218–229. https://doi.org/10.1080/24748668.2023.2280348

Guru, C. S., Datta, K., Krishnan, A., Kumar, A., Mahajan, U., Jhajharia, S., Kumar, M., Kumar, A., Yadav, N., & Singh, A. (2019). Sleep pattern and sports performance of adolescent archers: A pre–post-intervention pilot study. *Sleep and Vigilance, 3*(2), 143–149. https://doi.org/10.1007/s41782-019-00073-5

Guru, C. S., Krishnan, A., Mahajan, U., & Sharma, D. (2020). Heart rate values during shooting is a field-side performance analysis tool in archery – A study of elite Indian archers. *International Journal of Sport Studies for Health, 3*(1), Article e99687. https://doi.org/10.5812/intjssh.99687

Handayani, L., Kusnanik, N. W., Rusdiawan, A., Rasyid, M. L. S. A., García-Jiménez, J. V., & Pranoto, A. (2024). Do the physical elements improving archery performance? *Retos, 56,* 385–389. https://doi.org/10.47197/retos.v56.103791

Harvey, N. R., Voisin, S., Dunn, P. J., Sutherland, H., Yan, X., Jacques, M., Papadimitriou, I. D., Haseler, L. J., Ashton, K. J., Haupt, L. M., Eynon, N., & Griffiths, L. R. (2020). Genetic variants associated with exercise performance in both moderately trained and highly trained individuals. *Molecular Genetics and Genomics, 295*(2), 515–523. https://doi.org/10.1007/s00438-019-01639-8

Hoog Antink, C., Braczynski, A. K., & Ganse, B. (2021). Learning from machine learning: Prediction of age-related athletic performance decline trajectories. *GeroScience, 43*(5), 2547–2559. https://doi.org/10.1007/s11357-021-00411-4

Keast, D., & Elliott, B. (1990). Fine body movements and the cardiac cycle in archery. *Journal of Sports Sciences, 8*(3), 203–213. https://doi.org/10.1080/02640419008732146

Kim, H.-B., Kim, S.-H., & So, W.-Y. (2015). The relative importance of performance factors in Korean archery. *Journal of Strength and Conditioning Research, 29*(5), 1211–1219. https://doi.org/10.1519/JSC.0000000000000687

Kolayiş, İ. E., & Mimaroğlu, E. (2008). Okçuluk Milli Takımının antrenman ortamında kalp atım hızı ve nişan alma süresinin atış puanı üzerindeki etkileri [The effects of heart rate and aiming time on performance in Turkish national archery team]. *Journal of Human Sciences, 5*(1). https://j-humansciences.com/ojs/index.php/ijhs/article/view/363/276

Lee, K.-H. (2009). Evaluation of attention and relaxation levels of archers in shooting process using brain wave signal analysis algorithms evaluation of attention and relaxation levels of archers in shooting process using brain wave signal analysis algorithms. *Science of Emotion and Sensibility, 12*(3), 341–350. https://koreascience.kr/article/JAKO200935535844989.pdf

Lu, Y., & Zhong, S. (2023). Contactless real-time heart rate predicts the performance of elite athletes: Evidence from Tokyo 2020 Olympic archery competition. *Psychological Science, 34*(3), 384–393. https://doi.org/10.1177/09567976221143127

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 30). https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

Mahajan, U., Krishnan, A., Malhotra, V., Sharma, D., & Gore, S. (2019). Predicting fitness and performance of diving using machine learning algorithms. In *2019 IEEE Pune Section International Conference (PuneCon)* (pp. 1–5). https://doi.org/10.1109/PuneCon46936.2019.9105817

Mahajan, U., Krishnan, A., Malhotra, V., Sharma, D., & Gore, S. (2021). Machine learning feature selection in archery performance. In S. N. Merchant, K. Warhade, & D. Adhikari (Eds.), *Advances in signal and data processing* (pp. 561–573). Springer. https://doi.org/10.1007/978-981-15-8391-9_41

Marfell-Jones, M., Odds, T., Stewart, A., & Carter, L. (2006). *International standards for anthropometric assessment.* International Society for the Advancement of Kinanthropometry.

Mohamed, M. N., Wan Norman, W. M. N., Linoby, A., Sariman, M. H., & Mohd Azam, M. Z. (2014). Breathing pattern influence to the shooting performance. In R. Adnan, S. I. Ismail, & N. Sulaiman (Eds.), *Proceedings of the International Colloquium on Sports Science, Exercise, Engineering and Technology 2014 (ICoSSEET 2014)* (pp. 321–333). Springer Singapore.

Muazu Musa, R., Taha, Z., P. P. Abdul Majeed, A., & Abdullah, M. R. (2019a). Bio-physiological indicators in evaluating archery performance. In R. Muazu Musa, Z. Taha, A. P. P. Abdul Majeed, & M. R. Abdullah (Eds.), *Machine learning in sports: Identifying potential archers* (pp. 13–20). Springer Singapore. https://doi.org/10.1007/978-981-13-2592-2_2

Muazu Musa, R., Taha, Z., P. P. Abdul Majeed, A., & Abdullah, M. R. (2019b). Anthropometry correlation towards archery performance. In R. Muazu Musa, Z. Taha, A. P. P. Abdul Majeed, & M. R. Abdullah (Eds.), *Machine learning in sports: Identifying potential archers* (pp. 29–35). Springer Singapore. https://doi.org/10.1007/978-981-13-2592-2_4

Neumann, D. L., & Thomas, P. R. (2009). The relationship between skill level and patterns in cardiac and respiratory activity during golf putting. *International Journal of Psychophysiology, 72*(3), 276–282. https://doi.org/10.1016/j.ijpsycho.2009.01.001

Robazza, C., Bortoli, L., & Nougier, V. (1998). Physiological arousal and performance in elite archers: A field study. *European Psychologist, 3*(4), 263–270. https://doi.org/10.1027/1016-9040.3.4.263

Robazza, C., Bortoli, L., & Nougier, V. (1999). Emotions, heart rate and performance in archery. *Journal of Sports Medicine and Physical Fitness, 39*(2), 169–176. https://www.minervamedica.it/en/journals/sports-med-physical-fitness/article.php?cod=R40Y1999N02A0169

Rusdiawan, A., Kusuma, D. A., Rasyid, M. L. S. A., García-Jiménez, J. V., Purnomo, M., Wismanadi, H., Siantoro, G., Lani, A., Irmawati, F., & Fitriyah Ningsih, Y. (2024). Physical capacity and performance correlation in sub-elite Indonesian archery athletes. *Retos, 60,* 1093–1101. https://doi.org/10.47197/retos.v60.107017

Shaari, J. S., Nizam, N. F. D. M., Nor, M. A. M., Halim, N. A. A., Aiman, S., & Prasetyo, Y. (2023). Relationship between competitive anxiety and performance:Aa study of the archers at the national competition. In B. A. M et al. (Eds.), *Proceedings of the 8th ACPES (ASEAN Council of Physical Education and Sport) International Conference*. European Alliance for Innovation. https://doi.org/10.4108/eai.28-10-2022.2327500

Sellami, M., Elrayess, M. A., Puce, L., & Bragazzi, N. L. (2021). Molecular big data in sports sciences: State-of-art and future prospects of OMICS-based sports sciences. *Frontiers in Molecular Biosciences, 8,* Article 815410. https://doi.org/10.3389/fmolb.2021.815410

Smith, R. E., Smoll, F. L., & Schutz, R. W. (1990). Measurement and correlates of sport-specific cognitive and somatic trait anxiety: The Sport Anxiety Scale. *Anxiety Research, 2*(4), 263–280. https://doi.org/10.1080/08917779008248733

Söğüt, F., Şimşek, M., & Kesilmiş, İ. (2025). Predicting high potential archers by the quite eye duration parameter. *Sports Engineering, 28*(1), Article 9. https://doi.org/10.1007/s12283-025-00492-w

Soylu, A. R., Ertan, H., & Korkusuz, F. (2006). Archery performance level and repeatability of event-related EMG. *Human Movement Science, 25*(6), 767–774. https://doi.org/10.1080/14763141.2022.2049357

Taha, Z., Musa, R. M., Abdul Majeed, A. P., Abdullah, M. R., Ab. Nasir, A. F., & Hassan, M. H. A. (2018). Classification of high performance archers by means of bio-physiological performance variables via k-Nearest Neighbour classification model. In M. H. A. Hassan (Ed.), *Intelligent manufacturing & mechatronics. Lecture notes in mechanical engineering* (pp. 377–384). Springer Singapore. https://doi.org/10.1007/978-981-10-8788-2_33

Taha, Z., Musa, R. M., Abdul Majeed, A. P., Abdullah, M. R., & Hassan, M. H. A. (2018). Talent identification of potential archers through fitness and motor ability performance variables by means of artificial neural network. In M. H. A. Hassan (Ed.), *Intelligent manufacturing & mechatronics. Lecture notes in mechanical engineering* (pp. 371–376). Springer Singapore. https://doi.org/10.1007/978-981-10-8788-2_32

USA Archery. (Ed.) (2012). *Archery*. Human Kinetics.

Vendrame, E., Belluscio, V., Truppa, L., Rum, L., Lazich, A., Bergamini, E., & Mannini, A. (2024). Performance assessment in archery: A systematic review. *Sports Biomechanics, 23*(12), 2444–2446. https://doi.org/10.1080/14763141.2022.2049357

Wang, K., Li, Y., Liu, H., Zhang, T., & Luo, J. (2023). Relationship between pistol players' psychophysiological state and shot performance: Activation effect of EEG and HRV. *Scandinavian Journal of Medicine & Science in Sports, 33*(1), 84–98. https://doi.org/10.1111/sms.14253

Appendix

Table A1 Variable names, types, and descriptions

Name	Type	Description
Predictors/Independent variables		
group_NoneliteRecurve	Binary	Group – non-elite recurve (ref: non-elite compound)
group_XEliteCompound	Binary	Group – elite compound (ref: non-elite compound)
group_XEliteRecurve	Binary	Group – elite recurve (ref: non-elite compound)
hrb_5sec	Numeric	Heart rate at –5 s arrow release (beats per minute)
hrb_4sec	Numeric	Heart rate at –4 s arrow release (beats per minute)
hrb_3sec	Numeric	Heart rate at –3 s arrow release (beats per minute)
hrb_2sec	Numeric	Heart rate at –2 s arrow release (beats per minute)
hrb_1sec	Numeric	Heart rate at –1 s arrow release (beats per minute)
hr_0sec	Numeric	Heart rate at 0 s arrow release (beats per minute)
hra_1sec	Numeric	Heart rate at +1 s arrow release (beats per minute)
hra_2sec	Numeric	Heart rate at +2 s arrow release (beats per minute)
hra_3sec	Numeric	Heart rate at +3 s arrow release (beats per minute)
hra_4sec	Numeric	Heart rate at +4 s arrow release (beats per minute)
hra_5sec	Numeric	Heart rate at +5 s arrow release (beats per minute)
trgphase_Xprecompetition	Binary	Training phase – pre competition
age	Numeric	Age (years)
sportsage	Numeric	Sports age (years)
trgage	Numeric	Training age (years)
ht	Numeric	Height (cm)
wt	Numeric	Weight (kg)
whr	Numeric	Waist-to-hip ratio
fat	Numeric	Body fat (%)
rt_grpstr	Numeric	Right hand grip strength (kg)
lt_grpstr	Numeric	Left hand grip strength (kg)
resthr	Numeric	Resting heart rate (beats per minute)
restbp_sys	Numeric	Resting blood pressure – systolic (mmHg)
restbp_dia	Numeric	Resting blood pressure – diastolic (mmHg)
hr_bwup	Numeric	Heart rate before warm–up
somanx	Numeric	Somatic trait anxiety (9–36)
conganx	Numeric	Cognitive trait anxiety (7–28)
cdanx	Numeric	Concentration disruption trait anxiety (5–20)
sleequal_ln_Good	Binary	Sleep quality last night – good (ref: average)
high_score	Numeric	Highest score in previous competitions
highscr_compcat_International	Binary	Type of previous highest performing competitions – international (ref: below national)
highscr_compcat_National	Binary	Type of previous highest performing competitions – national (ref: below national)
Dependent/outcome variable		
sumscore	Numeric	Total archery score in Session 1, Session 2