

Research article

# Machine Learning–Based Classification of Alertness Levels in Elite Shooting Athletes Using Heart Rate Variability

Jiaojiao Lu<sup>1,2</sup>, Jun Qiu<sup>2</sup>✉ and Yan An<sup>2</sup>

<sup>1</sup> School of Exercise and Health, Shanghai University of Sport, Shanghai, China.

<sup>2</sup> Shanghai Research Institute of Sports Science (Shanghai Anti-Doping Agency), Shanghai, China

## Abstract

This study aims to develop a predictive model for alertness levels in elite shooting athletes by analyzing heart rate variability (HRV) dynamics under simulated competitive stress. 83 national-level shooting athletes completed a 60-minute Psychomotor Vigilance Task (PVT) protocol designed to mimic the sustained attentional demands of a competition, while HRV data were continuously recorded. Pearson correlation analysis identified HRV features associated with behavioral performance. Key predictors were selected via recursive feature elimination with Random Forest. Four machine learning algorithms—Support Vector Machine (SVM), Random Forest (RF), XGBoost, and AdaBoost—were employed to construct classification models for alertness. Model performance was evaluated using accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC). SHAP analysis was applied to interpret feature contributions. The binary classification framework (optimal vs. sub-optimal alertness) demonstrated superior reliability over multi-class approaches. The AdaBoost model achieved the best performance, with an accuracy of 0.75, an F1-score of 0.73, and an AUC of 0.77. SHAP analysis revealed that the very low frequency percentage (VLF%) was the most critical predictor, followed by the SD2/SD1 ratio. Notably, elevated VLF% values were associated with lower alertness levels. The binary classification model, integrating key HRV indices (notably VLF%) with the AdaBoost algorithm, can effectively distinguish alertness levels in shooting athletes during simulated competitive stress. This approach provides a validated, non-invasive tool for objective psychophysiological monitoring in training, offering actionable insights for pre-competition readiness assessment.

**Key words:** Heart Rate Variability, Shooters, Vigilance, Machine Learning, Psychomotor Vigilance Task.

## Introduction

In competitive precision sports such as shooting, athletes are required to maintain an exceptionally high and stable level of psychophysiological readiness (Chang et al., 2020), often referred to as vigilance. This state encompasses the capacity to sustain focused attention and optimal responsiveness over extended periods while effectively inhibiting distractions (Reifman et al., 2018). However, high-stakes competitive stress can disrupt this delicate state, leading to specific impairments such as increased aiming fluctuation, delayed trigger control, and a consequent decline in performance accuracy (Diaz et al., 2013; Hashemi et al., 2019). Therefore, maintaining optimal alertness—defined here as the immediate, tonic level of central nervous system preparedness—is a critical determinant of competitive success (Mah et al., 2011; Torres and Kim, 2019).

Despite its importance, the real-time assessment of

vigilance remains a challenge in field settings. Coaches often rely on subjective rating scales (e.g., Multidimensional Fatigue Inventory-20 (MFI-20), Visual Analog Scale (VAS), Stanford Sleepiness Scale (SSS)) (Mah et al., 2011; Munguía-Izquierdo et al., 2012; Tan et al., 2023), which are prone to bias and lack continuity. While the Psychomotor Vigilance Task (PVT) is considered the gold standard for quantifying vigilance decrement (Sun et al., 2024), its requirement for active participant responses interrupts training flow. Similarly, neurophysiological measures like EEG and eye-tracking, though precise, require sophisticated equipment and strict control, limiting their utility for capturing real-time fluctuations during actual competition (Laborde et al., 2017; Li et al., 2023). Consequently, there is a critical need for an objective, non-invasive, and unobtrusive methodology for monitoring athletes psychophysiological states.

Heart rate variability (HRV) offers a promising solution. As a non-invasive metric derived from ECG, HRV quantifies the autonomic nervous system's modulation, which is intrinsically linked to psychophysiological arousal and vigilance (Laborde et al., 2017; Li et al., 2023). Its compatibility with wearable devices makes it suitable for ambulatory monitoring in sports (Li et al., 2016; Vitale et al., 2019). To leverage this data, machine learning (ML) techniques have been increasingly applied to model the complex, non-linear relationships between HRV indices and vigilance levels (Balakarthikeyan et al., 2023; Li et al., 2023; Ma et al., 2024). For instance, Zhou and Zhang, demonstrated the feasibility of using Support Vector Machines (SVM) to classify vigilance (Zhou and Zhang, 2022). However, the accuracy and generalizability of such models depend heavily on the ecological validity of the training data. Previous studies have often relied on short-duration laboratory tests or non-elite populations, which may not authentically reflect the state fluctuations experienced by professional athletes during competition.

The primary objective of this study was to develop and validate a machine learning–based classification model that discriminates between optimal and sub-optimal alertness states in elite shooting athletes using dynamic HRV features recorded during a prolonged, ecologically valid vigilance task. The secondary objective was to identify the most informative HRV predictors via SHAP analysis, thereby informing the physiological basis of the model. We hypothesized that HRV features recorded during the 60-minute simulated competition task would provide sufficient discriminative information to classify alertness levels with an AUC exceeding 0.70, and that

frequency-domain indices—particularly those reflecting slower autonomic oscillations—would emerge as the most critical predictors.

## Methods

### Sample size calculation

The state of “sub-optimal” alertness was operationally defined based on behavioral performance in the Psychomotor Vigilance Task (PVT). A PVT trial was classified as representing a sub-optimal alertness state if the participant's reaction time exceeded 500 ms (Basner and Dinges, 2011). The operational threshold of  $RT > 500$  ms to define a sub-optimal alertness epoch was adopted in accordance with the well-established PVT convention, wherein a response exceeding 500 ms is classified as an attentional lapse (Dinges and Powell, 1985; Reifman et al., 2018; Van Dongen et al., 2003). This threshold reflects a response substantially slower than the typical alert reaction time of 200 - 300 ms and ensures comparability with the broader PVT literature.

Preliminary data from an internal pilot study ( $n \approx 20$  athletes) conducted by our research group indicated an expected model discrimination performance (Area Under the Curve, AUC) of 0.75 for identifying this state, with an estimated positive event rate of 35%. The required sample size was calculated using Obuchowski's method for correlated data ( $\alpha = 0.05$ , power = 80%), which determined that a minimum of 54 positive event epochs (i.e., PVT trials with  $RT > 500$  ms) were needed.

From the final dataset, 480 valid PVT epochs, with each epoch defined as a 10-minute task block and meeting the data quality criteria, were obtained. Within this dataset, 168 epochs were identified as positive cases (sub-optimal alertness), which significantly exceeded the minimum requirement calculated a priori. Post-hoc evaluation confirmed that the study achieved robust statistical power. Given that the a priori power analysis required a minimum of 54 positive events to detect an AUC of 0.75 with 80% power, our final dataset—comprising 168 positive events and achieving an observed AUC of 0.77—substantially exceeded these requirements. Therefore, the observed power for the primary classification metric (AUC) is inherently greater than the predefined 80% threshold.

### Participants

A total of 83 elite shooting athletes from Shanghai (48 males, 35 females; mean age  $18.4 \pm 2.5$  years; mean body mass  $62.3 \pm 9.1$  kg; mean height  $168.7 \pm 7.4$  cm) were enrolled, all certified as national first-grade athletes or above. In terms of competitive level, 9 athletes held International Master of Sport certification, 25 held National Master of Sport certification, and the remaining 49 were certified national first-grade athletes. Athletes specialized in 10-meter air rifle ( $n = 47$ ) or 10-meter air pistol ( $n = 36$ ) events, with mean training experience of  $5.8 \pm 2.3$  years and a weekly training volume of  $28.6 \pm 4.2$  hours, conducted predominantly in the morning as part of the team's standardized schedule.

The study was approved by the Ethics Committee of the Shanghai Research Institute of Sports Science

(Shanghai Anti-Doping Agency) [LLSC20250005], All participants were informed about the study protocol and provided written informed consent to participate in the study. All procedures were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

### Inclusion and exclusion criteria

Participants were eligible for inclusion if they met the following criteria: (i) certified elite shooting athletes at or above the national first-grade level; (ii) actively engaged in regular training programs; and (iii) able to complete the full experimental protocol.

Participants were excluded if they met any of the following criteria: (i) self-reported chronic sleep disorders; (ii) acute illness at the time of testing; or (iii) inability to complete physiological or behavioral measurements.

To control for potential confounding factors, several pre-test restrictions were applied. All participants were non-smokers and were required to abstain from alcohol consumption for at least 24 hours prior to testing, in accordance with standard athlete health management protocols. Participants were further instructed to refrain from caffeine-containing beverages, stimulants, and ergogenic supplements for at least 12 hours before testing to minimize their influence on autonomic nervous system activity.

Sleep-related factors were partially controlled by excluding participants with self-reported chronic sleep disorders; however, formal objective assessment of baseline sleep quality (e.g., PSQI or actigraphy) was not conducted prior to testing.

Female participants ( $n = 35$ ) were included in the study; however, menstrual cycle phase was not systematically recorded or controlled, and testing was not restricted to specific cycle phases.

All experimental sessions were conducted in the morning prior to routine training to reduce variability associated with diurnal fluctuations in physiological and cognitive performance.

### Data collection

All testing sessions were conducted between 08:00 and 09:30 in the morning, prior to the commencement of daily training. This timing was standardized to control for diurnal variation in both HRV and vigilance performance, given that HRV exhibits well-documented circadian rhythmicity in autonomic tone (Boudreau et al., 2013; Vitale et al., 2019) and alertness levels fluctuate in accordance with circadian phase (Laborde et al., 2017). As HRV monitoring is a routine component of the team's athlete health management program, consistent morning testing prior to training is an established practice.

### Alertness task

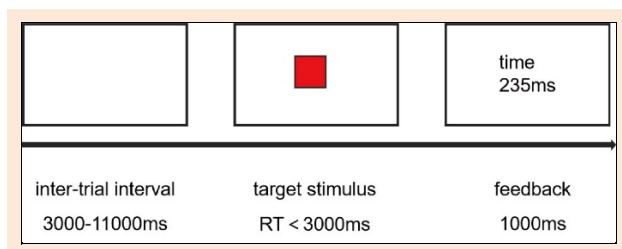
To quantify vigilance levels under conditions simulating the sustained attention demands of precision shooting, the PVT was employed. The task was programmed and presented using PsychoPy (v2022.2.5) (Peirce et al., 2019), a free, open-source experiment builder widely used in behavioral research for its millisecond-accurate stimulus

delivery and response recording. The PVT was selected as the alertness assessment tool because it is the gold-standard validated paradigm for measuring sustained attention and detecting vigilance decrements (Basner and Dinges, 2011; Dinges and Powell, 1985). Its simple reaction-time format, freedom from learning effects, and sensitivity to alertness-reducing factors make it ideal for use in athletic populations (Reifman et al., 2018), and its application in sport science contexts has been established in prior research (Tan et al., 2023; Xie and Ma, 2025).

At the start of the experiment, participants were instructed to rest their dominant hand on a keyboard spacebar and to maintain fixation on a central point on the screen. Each trial began with a blank-screen inter-trial interval varying randomly between 2 and 10 seconds, followed by the appearance of a visual stimulus (a millisecond timer starting from 0). Participants were required to press the spacebar as quickly as possible upon stimulus onset. Immediately after each response, the reaction time (RT) in milliseconds was displayed for 1 second, followed by the next trial. If no response was detected within a pre-defined window, the trial terminated automatically after a timeout, and the next trial began.

To ecologically mirror the temporal structure of 10-meter air rifle and pistol events—in which athletes typically complete their match within approximately 60 minutes amid varying pacing strategies—the PVT was structured into six consecutive 10-minute blocks, totaling 60 minutes of testing. Each 10-minute block contained 80 trials. This block duration was designed to simulate the sustained focus required during a typical shooting series, where athletes repeatedly engage in aiming, breath control, and trigger execution over extended periods without prolonged breaks. Throughout the entire task, HRV was recorded continuously.

A valid response was defined as a keyboard press occurring after stimulus onset and within the response window. Performance metrics derived from the PVT included: number of valid responses, mean RT, median RT, mean reciprocal RT (1/RT), the average of the fastest 10% of RTs, and the average of the slowest 10% of RTs. These indices provide a multi-faceted assessment of vigilance, with reciprocal RT and fastest RTs reflecting optimal alertness, and slowest RTs capturing lapses in attention. A one-way analysis of variance (ANOVA) was conducted to examine differences in reaction times across experimental conditions or groups. Alertness task is presented in Figure 1.



**Figure 1.** Schematic diagram of alertness task.

**Electrocardiographic signal acquisition and processing**  
Electrocardiographic (ECG) data continuously recorded

throughout the vigilance task using a Polar V800 heart rate monitor (Polar Electro Oy, Finland) with a sampling frequency of 1000 Hz. The raw ECG signal was first band-pass filtered (0.5 - 35 Hz) to attenuate baseline wander and high-frequency noise. Subsequently, the derived RR interval time series were processed to correct for artifacts and ectopic beats, which are common in ambulatory recordings. This correction was performed using the built-in "Artifact Correction Algorithm" within Kubios HRV Premium software (version 3.5). The software's automatic correction mode, with its threshold set to "Medium", was applied to identify and interpolate spurious or missing beats, ensuring the integrity of the inter-beat interval data for subsequent analysis. Under this 'Medium' correction threshold, the proportion of corrected RR intervals was consistently below 2% per participant per epoch. This low correction rate falls well within the accepted quality threshold for HRV analysis (Laborde et al., 2017; Malik et al., 1996), thereby minimizing any potential distortion of frequency-domain and nonlinear HRV metrics and indicating high signal integrity across the dataset. From the artifact-corrected RR interval series, standard frequency-domain HRV parameters were computed using Kubios. The extracted indices included: very low frequency (VLF: 0.01 - 0.04 Hz), low frequency (LF: 0.04 - 0.15 Hz), high frequency (HF: 0.15 - 0.40 Hz), the LF/HF ratio, and total power. These parameters provide a comprehensive quantification of autonomic nervous system activity related to the vigilance state.

### Alertness level labeling

To ensure the robustness of the predictive model, alertness levels were operationally defined based on the distribution of mean reaction times (RT) for each 10-minute block. Initially, a fine-grained five-level classification was established using a standard deviation (SD) deviation method based on the entire dataset's distribution. The specific cut-off values and data distribution for these five levels are detailed in Supplementary Table 1. However, for practical field applications, distinguishing between broad alertness states is often more meaningful than detailed categorization. Therefore, the five levels were consolidated into two simplified frameworks:

1. Three-class framework: Classifying states into High, Moderate, and Low alertness (Supplementary Table 2).
2. Binary-class framework: Distinguishing simply between Optimal (Top 40%) and Sub-optimal (Bottom 40%) alertness (Supplementary Table 3).

The binary framework was prioritized for the final model development to maximize discriminative power and ensure a balanced dataset for machine learning training.

### Machine learning and performance evaluation

Machine learning and performance evaluation were conducted by first splitting the original dataset into a training dataset and a test dataset at a ratio of 7:3. During model development, feature selection was performed using random forest (RF) combined with recursive feature elimination to identify relevant HRV features, including time-domain, frequency-domain, nonlinear, and autonomic nervous system indices. The selected features were then

used to build vigilance level prediction models via four machine learning algorithms: support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost). Hyperparameter tuning was carried out using grid search with five-fold cross-validation. Hyperparameter tuning was conducted using exhaustive grid search combined with five-fold cross-validation (folds defined at the subject level to prevent data leakage). The search grids were as follows: for SVM, regularization parameter  $C \in \{0.1, 1, 10, 100\}$  and kernel coefficient  $\gamma \in \{\text{'scale'}, \text{'auto'}, 0.001, 0.01\}$ ; for RF, number of estimators  $\in \{50, 100, 200\}$  and maximum tree depth  $\in \{\text{None}, 5, 10, 20\}$ ; for XGBoost, number of estimators  $\in \{50, 100, 200\}$ , learning rate  $\in \{0.01, 0.1, 0.3\}$ , and maximum depth  $\in \{3, 5, 7\}$ ; for AdaBoost, number of estimators  $\in \{50, 100, 200\}$  and learning rate  $\in \{0.01, 0.1, 0.5, 1.0\}$ . AUC was used as the optimization metric throughout. Model performance was evaluated on an independent test set using accuracy, specificity, sensitivity, F1 score, and AUC as evaluation metrics. Furthermore, feature importance ranking and SHAP (Shapley Additive Explanations) analysis were applied to identify key factors associated with vigilance levels.

### Statistical analysis

Data preprocessing and statistical analyses were conducted using JASP (Version 0.19.3). Prior to analysis, the normality of continuous variables was assessed using the Shapiro–Wilk test. Descriptive statistics are reported as  $M \pm SD$ . A one-way ANOVA was performed to compare reaction time differences across the six PVT blocks. To feature selection, Pearson correlation analysis was employed to evaluate the linear associations between all extracted HRV indices (time-domain, frequency-domain, and nonlinear) and behavioral alertness metrics. A

correlation was considered statistically significant at  $p < 0.05$ . Features showing significant correlations were identified as candidate inputs for machine learning models.

## Results

### PVT test

The results of the six consecutive PVT blocks are presented in Table 1. The mean reaction time, median reaction time, and the slowest 10% reaction time generally increased across blocks, while the fastest 10% reaction time showed a decreasing trend. ANOVA revealed statistically significant differences in the mean reaction time, median reaction time, and slowest 10% reaction time across the six PVT blocks. However, no significant difference was observed for the fastest 10% reaction time ( $F = 0.88, p > 0.05$ ).

### Dynamic Changes in Heart Rate Variability

HRV parameters exhibited distinct temporal patterns throughout the task, reflecting a shift in autonomic regulation.

### Time-domain and Frequency-domain Analysis

Time-domain analysis (Table 2) revealed a gradual increase in overall variability, with SDNN rising from  $44.3 \pm 17.62$  ms (Block 1) to  $47.51 \pm 16.87$  ms (Block 6). In the frequency domain (Table 3), Total Power exhibited an increasing trend. While the absolute power of VLF (aVLF) and LF (aLF) increased significantly over time, the normalized values (LF norm, HF norm) and the LF/HF ratio remained relatively stable. This suggests that while total autonomic modulation increased with time-on-task, the sympathovagal balance did not show a drastic linear shift in the frequency domain.

**Table 1.** Changes in reaction time metrics across six consecutive PVT blocks ( $M \pm SD$ )

Indicators	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	F
Mean reaction time (s)	272.99 $\pm$ 35.83	286.2 $\pm$ 42.86	281.65 $\pm$ 47.39	290.08 $\pm$ 49.53	291.39 $\pm$ 57.72	297.8 $\pm$ 66.53	3.45**
Median reaction time(s)	261.25 $\pm$ 30.54	272.87 $\pm$ 36.22	269.7 $\pm$ 45.35	275.32 $\pm$ 44.45	275.51 $\pm$ 53.7	281.53 $\pm$ 56.98	3.06*
Fastest 10% reaction time(s)	203.55 $\pm$ 51.96	202.78 $\pm$ 63.49	192.91 $\pm$ 75.83	194.52 $\pm$ 75.24	185.47 $\pm$ 88.35	191.19 $\pm$ 86.54	0.88
Slowest 10% reaction time(s)	397.12 $\pm$ 100.75	441.17 $\pm$ 142.83	434.98 $\pm$ 122.74	462.77 $\pm$ 153.93	477.97 $\pm$ 136.45	492.49 $\pm$ 169.4	7.51***

\*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .

**Table 2.** Changes in time-domain features across six consecutive PVT blocks ( $\bar{x} \pm s$ ).

Time-domain features	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Mean RR	779.82 $\pm$ 105.37	775.4 $\pm$ 98.24	774.14 $\pm$ 94.6	779.91 $\pm$ 92.84	780.01 $\pm$ 91.09	784.21 $\pm$ 90.75
SDNN	44.3 $\pm$ 17.62	45.06 $\pm$ 18.31	46.42 $\pm$ 19.57	46.38 $\pm$ 15.26	49.11 $\pm$ 18.92	47.51 $\pm$ 16.87
Mean HR	78.28 $\pm$ 10.16	78.53 $\pm$ 9.31	78.59 $\pm$ 9.07	77.96 $\pm$ 8.82	77.95 $\pm$ 8.95	77.49 $\pm$ 8.67
SDHR	4.58 $\pm$ 2.45	4.73 $\pm$ 2.6	4.95 $\pm$ 2.86	4.76 $\pm$ 1.78	5.16 $\pm$ 2.58	4.8 $\pm$ 2.02
HR Max-min	27.21 $\pm$ 8.25	28.9 $\pm$ 9.65	29.02 $\pm$ 8.1	29.65 $\pm$ 8.1	31.88 $\pm$ 13.58	30.21 $\pm$ 7.37
RMSSD	40.26 $\pm$ 23.77	38.76 $\pm$ 27.19	39.63 $\pm$ 28.49	39.16 $\pm$ 19.71	41.79 $\pm$ 27.43	40.07 $\pm$ 23.13
NN50	96.19 $\pm$ 81.65	81.41 $\pm$ 65.68	81.31 $\pm$ 67.08	84.62 $\pm$ 66.54	83.82 $\pm$ 68.01	82.37 $\pm$ 61.77
pNN50	18.91 $\pm$ 16.54	15.67 $\pm$ 13.58	15.6 $\pm$ 13.8	16.3 $\pm$ 13.76	16.12 $\pm$ 14.05	16.2 $\pm$ 13.08
RR tri index	11.8 $\pm$ 3.86	11.56 $\pm$ 3.39	12.14 $\pm$ 3.82	12.12 $\pm$ 3.44	12.28 $\pm$ 3.59	11.96 $\pm$ 3.22
TINN	234.24 $\pm$ 102.34	244.75 $\pm$ 125.05	245.86 $\pm$ 123.6	250.46 $\pm$ 104.4	272.04 $\pm$ 148.74	254.37 $\pm$ 132.94

### Nonlinear feature changes

Table 4 presents the changes in nonlinear features across six consecutive PVT blocks. As shown in the table, alpha 1, alpha 2 and SD1 exhibited minimal variation and remained relatively stable throughout the task. In contrast, SD2 and the SD2/SD1 ratio gradually increased, while ApEn and SampEn showed a decreasing trend. Specifically, the mean SD2 increased from  $55.33 \pm 19.75$  to  $60.34 \pm 19.17$ , and the SD2/SD1 ratio rose from  $2.16 \pm 0.56$  to  $2.35 \pm 0.59$ . Meanwhile, ApEn decreased from  $1.30 \pm 0.07$  to  $1.27 \pm 0.10$ , and SampEn declined from  $1.69 \pm 0.20$  to  $1.61 \pm 0.24$ .

### Correlation analysis

Pearson correlation analysis was conducted to examine the associations between HRV indices and PVT reaction times. As shown in Table 5, all time-domain and frequency-domain HRV metrics, except for peakHF, demonstrated certain degrees of correlation with PVT reaction time.

### Alertness level classification

The three-class model results revealed a consistent pattern across all four algorithms: the 'Moderate alertness' class was poorly identified, with Recall values ranging from only 0.02 (SVM, AdaBoost) to 0.15 (XGBoost) and F1-scores between 0.05 and 0.22 (Table 6). Classification error analysis showed that misclassified 'Moderate' epochs were overwhelmingly assigned to the adjacent High or Low classes, indicating severe boundary confusion. In contrast, the binary classification framework eliminated this ambiguous intermediate zone, significantly improving the optimal AdaBoost model's AUC from 0.67 (in the three-class framework) to 0.77 (see Table 6 and Table 7 for comparison).

Based on the binary classification criteria established in the Methods section (Optimal: Top 40% vs. Sub-optimal: Bottom 40%), the final dataset was perfectly balanced. As shown in the distribution histogram (Figure 2), the dataset comprised 240 epochs labeled as "High Alertness" (mean RT < 273.98 ms) and 240 epochs labeled

as "Low Alertness" (mean RT  $\geq 319.60$  ms). This balanced class distribution (1:1 ratio) facilitates the training of unbiased machine learning models without the need for synthetic oversampling techniques.

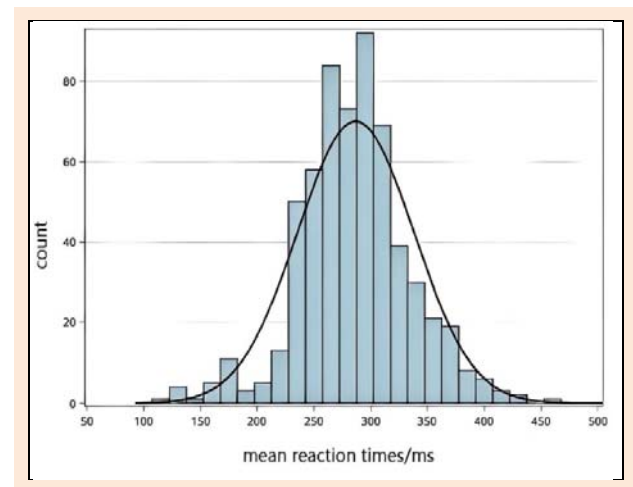


Figure 2. Distribution of mean reaction time data.

### Model development and evaluation

Using the selected HRV features, four machine learning algorithms were evaluated. Table 6 details the performance metrics. In the three-class framework, the models achieved moderate performance, with the Random Forest (RF) model yielding the highest accuracy of 0.57. However, transitioning to the binary classification framework significantly enhanced discriminative power. As detailed in Table 7, all four models showed improved metrics, with AdaBoost demonstrating superior performance, achieving the highest Accuracy (0.75), F1-score (0.73), and AUC (0.77). The ROC curves for the multi-class and binary frameworks are presented in Figure 3 and Figure 4, respectively, visually confirming the robust classification capability of the AdaBoost algorithm in distinguishing between optimal and sub-optimal alertness states.

Table 3. Changes in frequency-domain features across six consecutive PVT blocks ( $\bar{x} \pm s$ )

Frequency-domain features	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Total power	2051.4 $\pm$ 1632.3	2004.2 $\pm$ 1422.2	2193.6 $\pm$ 1620.0	2197.9 $\pm$ 1503.4	2416.0 $\pm$ 1826.2	2253.6 $\pm$ 1583.0
peakVLF	0.04 $\pm$ 0	0.04 $\pm$ 0	0.03 $\pm$ 0	0.04 $\pm$ 0	0.04 $\pm$ 0	0.04 $\pm$ 0
peakLF	0.09 $\pm$ 0.02	0.08 $\pm$ 0.02	0.08 $\pm$ 0.02	0.08 $\pm$ 0.02	0.08 $\pm$ 0.02	0.08 $\pm$ 0.02
peakHF	0.23 $\pm$ 0.07	0.23 $\pm$ 0.08	0.22 $\pm$ 0.07	0.22 $\pm$ 0.07	0.22 $\pm$ 0.07	0.22 $\pm$ 0.07
aVLF	123.56 $\pm$ 111.66	143.77 $\pm$ 104.6	172.25 $\pm$ 142.8	170.81 $\pm$ 124.67	186.22 $\pm$ 143.83	164.45 $\pm$ 108.56
aLF	1215.7 $\pm$ 899.1	1276.4 $\pm$ 863.7	1395.4 $\pm$ 905.4	1409.0 $\pm$ 914.5	1512.0 $\pm$ 982.3	1493.6 $\pm$ 1126.1
aHF	710.96 $\pm$ 932.5	583.0 $\pm$ 690.3	624.6 $\pm$ 827.3	616.6 $\pm$ 661.1	716.1 $\pm$ 938.5	594.6 $\pm$ 592.34
VLF(log)	4.48 $\pm$ 0.83	4.71 $\pm$ 0.76	4.88 $\pm$ 0.73	4.92 $\pm$ 0.68	4.97 $\pm$ 0.72	4.88 $\pm$ 0.7
LF(log)	6.84 $\pm$ 0.77	6.94 $\pm$ 0.69	7.01 $\pm$ 0.74	7.05 $\pm$ 0.66	7.12 $\pm$ 0.66	7.05 $\pm$ 0.74
HF(log)	6.11 $\pm$ 0.97	5.97 $\pm$ 0.87	6.01 $\pm$ 0.89	6.05 $\pm$ 0.86	6.08 $\pm$ 0.96	6.04 $\pm$ 0.84
VLF	7.25 $\pm$ 5.03	8.11 $\pm$ 4.79	9.11 $\pm$ 5.38	8.43 $\pm$ 4.03	8.96 $\pm$ 5.25	8.49 $\pm$ 4.5
LF	60.91 $\pm$ 13.5	64.83 $\pm$ 12.35	65.01 $\pm$ 11.28	65.09 $\pm$ 12.15	65.63 $\pm$ 12.14	65.29 $\pm$ 12.78
HF	31.79 $\pm$ 15.29	27.01 $\pm$ 12.68	25.83 $\pm$ 12	26.42 $\pm$ 13.16	25.36 $\pm$ 12.76	26.17 $\pm$ 12.04
LF norm	66.03 $\pm$ 15.42	70.67 $\pm$ 13.33	71.72 $\pm$ 12.53	71.29 $\pm$ 13.78	72.26 $\pm$ 13.29	71.31 $\pm$ 13.23
HF norm	33.92 $\pm$ 15.4	29.28 $\pm$ 13.31	28.22 $\pm$ 12.48	28.65 $\pm$ 13.73	27.68 $\pm$ 13.25	28.65 $\pm$ 13.21
LF/HF	2.6 $\pm$ 1.64	3.29 $\pm$ 2.37	3.22 $\pm$ 1.86	3.43 $\pm$ 2.41	3.44 $\pm$ 2.1	3.47 $\pm$ 2.59

**Table 4. Changes in nonlinear features across six consecutive PVT blocks ( $\bar{x} \pm s$ )**

Nonlinear feature	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
DFA alpha 1	1.18 ± 0.22	1.26 ± 0.2	1.26 ± 0.2	1.25 ± 0.21	1.26 ± 0.22	1.26 ± 0.21
DFA alpha 2	0.4 ± 0.12	0.42 ± 0.1	0.43 ± 0.11	0.43 ± 0.09	0.44 ± 0.11	0.43 ± 0.11
SD1	28.27 ± 16.97	27.08 ± 19.44	27.68 ± 20.37	27.19 ± 14.31	29.02 ± 19.73	27.75 ± 16.73
SD2	55.33 ± 19.75	56.95 ± 19.11	58.79 ± 20.66	59.02 ± 17.99	62.23 ± 20.41	60.34 ± 19.17
SD2/SD1	2.16 ± 0.56	2.32 ± 0.57	2.36 ± 0.55	2.34 ± 0.61	2.38 ± 0.58	2.35 ± 0.59
ApEn	1.3 ± 0.07	1.29 ± 0.09	1.28 ± 0.09	1.28 ± 0.09	1.27 ± 0.1	1.27 ± 0.1
SampEn	1.69 ± 0.2	1.63 ± 0.24	1.61 ± 0.24	1.61 ± 0.24	1.57 ± 0.25	1.61 ± 0.24

**Table 5. Correlation analysis between electrocardiogram metrics and PVT performance.**

Features	Mean reaction time		Median reaction time		Fastest 10% reaction time		Slowest 10% reaction time	
	r	p	r	p	r	p	r	p
Mean RR	0.054	0.190	0.054	0.190	0.081	0.048*	-0.003	0.936
SDNN	-0.068	0.096	-0.177	<.0001***	-0.237	<.0001***	-0.041	0.312
Mean HR	-0.197	<.0001***	-0.068	0.096	-0.099	0.016*	0.001	0.986
SDHR	-0.217	<.0001***	-0.197	<.0001***	-0.299	<.0001***	-0.020	0.620
HR Max-min	-0.155	<.0001***	-0.133	0.001**	-0.245	<.0001***	0.025	0.540
RMSSD	-0.139	<.0001***	-0.217	<.0001***	-0.275	<.0001***	-0.061	0.137
NN50	-0.120	0.003**	-0.155	<.0001***	-0.164	<.0001***	-0.064	0.120
pNN50	-0.228	<.0001***	-0.139	<.0001***	-0.137	0.001**	-0.067	0.100
RR tri index	-0.133	0.001**	-0.120	0.003**	-0.122	0.003**	-0.076	0.064
TINN	0.054	0.190	-0.228	<.0001***	-0.296	<.0001***	-0.031	0.449
Total power	0.025	0.541	-0.160	<.0001***	-0.208	<.0001***	-0.034	0.399
peakVLF	-0.108	0.008**	0.025	0.541	-0.002	0.960	0.004	0.924
peakLF	-0.061	0.133	-0.108	0.008**	0.058	0.156	-0.168	<.0001***
peakHF	0.036	0.384	-0.061	0.133	-0.031	0.451	-0.008	0.848
aVLF	-0.109	0.008**	0.036	0.384	-0.124	0.002**	0.130	0.001**
aLF	-0.198	<.0001***	-0.109	0.008**	-0.137	0.001**	-0.022	0.590
aHF	0.068	0.094	-0.198	<.0001***	-0.237	<.0001***	-0.064	0.116
VLF(log)	-0.111	0.007**	0.068	0.094	-0.121	0.003**	0.174	<.0001***
LF(log)	-0.194	<.0001***	-0.111	0.007**	-0.111	0.006**	-0.030	0.459
HF(log)	0.252	<.0001***	-0.194	<.0001***	-0.182	<.0001***	-0.071	0.082*
VLF	0.082	0.045*	0.252	<.0001***	0.057	0.160	0.234	<.0001***
LF	-0.170	<.0001***	0.082	0.045*	0.111	0.006**	0.007	0.867
HF	0.146	<.0001***	-0.170	<.0001***	-0.125	0.002**	-0.093	0.022*
LF norm	-0.145	<.0001***	0.146	<.0001***	0.121	0.003**	0.070	0.088
HF norm	-0.160	<.0001***	-0.145	<.0001***	-0.120	0.003**	-0.070	0.086
LF/HF	0.100	0.014*	0.100	0.014*	0.115	0.005**	0.027	0.515
DFA alpha 1	0.172	<.0001***	0.172	<.0001***	0.169	<.0001***	0.052	0.202
DFA alpha 2	0.164	<.0001***	0.164	<.0001***	-0.046	0.256	0.227	<.0001***
SD1	-0.234	<.0001***	-0.234	<.0001***	-0.280	<.0001***	-0.069	0.092
SD2	-0.141	<.0001***	-0.141	<.0001***	-0.195	<.0001***	-0.028	0.500
SD2/SD1	0.147	<.001***	0.148	<.0001***	0.136	0.001**	0.061	0.137
ApEn	0.122	0.003**	0.122	0.003**	0.184	<.0001***	-0.007	0.862
SampEn	0.097	0.017*	0.097	0.017*	0.198	<.0001***	-0.039	0.346

\* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001.

**Table 6. Performance of the alertness prediction models.**

Models	Three categories of alertness levels					Two categories of alertness levels				
	Precision	Recall	F1 score	Accuracy	AUC	Precision	Recall	F1 score	Accuracy	AUC
SVM	0.64	0.53	0.47	0.53	0.62	0.72	0.73	0.72	0.73	0.65
RF	0.58	0.57	0.53	0.57	0.66	0.74	0.74	0.71	0.74	0.72
XGBoost	0.52	0.52	0.49	0.52	0.65	0.73	0.74	0.73	0.74	0.72
Adaboost	0.49	0.53	0.47	0.53	0.67	0.74	0.75	0.73	0.75	0.77

**Table 7. Performance of four machine learning algorithms in the binary classification framework.**

Model	Accuracy	Precision (PPV)	Recall (Sensitivity)	Specificity	NPV	F1	AUC	Model
SVM	0.73	0.72	0.73	0.73	0.73	0.72	0.65	0.73
RF	0.74	0.74	0.74	0.74	0.74	0.71	0.72	0.74
XGBoost	0.74	0.73	0.74	0.74	0.74	0.73	0.72	0.74
Adaboost	0.75	0.74	0.75	0.75	0.75	0.73	0.77	0.75

PPV: Positive Predictive Value; NPV: Negative Predictive Value. Under the balanced test set (1:1 class ratio), PPV is equivalent to Precision, and Specificity is mathematically equivalent to NPV in this specific context; NPV values were derived from the confusion matrices of the independent test set.

### Model interpretability

To interpret the contribution of individual HRV features to the AdaBoost model, a SHAP analysis was performed. As shown in the feature importance ranking (Figure 5), VLF (%) was the most influential predictor, followed by the SD2/SD1 ratio and DFA alpha2, whereas TINN and HF (%) contributed minimally to the model output. The SHAP summary plot (Figure 6) further illustrates the directionality of feature effects. Higher values of VLF (%), SD2/SD1 ratio, alpha2, HR Max–min, and SampEn were associated with an increased likelihood of predicting low alertness (positive SHAP values). In contrast, higher Mean HR, HF power, TINN, and HF (%) were associated with predictions of high alertness.

### Discussion

This study presents a novel quantitative approach for assessing alertness in elite shooting athletes by integrating dynamic HRV monitoring with machine learning algorithms. The present findings broadly support our a priori hypothesis. The binary AdaBoost model achieved an

AUC of 0.77, exceeding the hypothesized threshold of 0.70, confirming that HRV features recorded during the 60-minute simulated competition task provide sufficient discriminative information for alertness classification in this elite population. Furthermore, consistent with our hypothesis, the frequency-domain index VLF% emerged as the most critical predictor in the SHAP analysis, underscoring the role of slower autonomic oscillations in encoding alertness states. However, the hypothesis regarding cross-demographic generalizability remains to be tested, given the single-sport, mixed-sex, and relatively young sample. Physiologically, shooting is a psychomotor task requiring intense top-down cognitive control and emotional regulation with minimal metabolic demand (Shao et al., 2020). Our findings support the neurovisceral integration model, where higher HRV—reflecting robust vagal tone—is associated with flexible autonomic adjustment and efficient cognitive resource allocation. This autonomic flexibility facilitates the suppression of task-irrelevant distractions, thereby supporting the sustained focus required for precision aim-and-trigger execution.

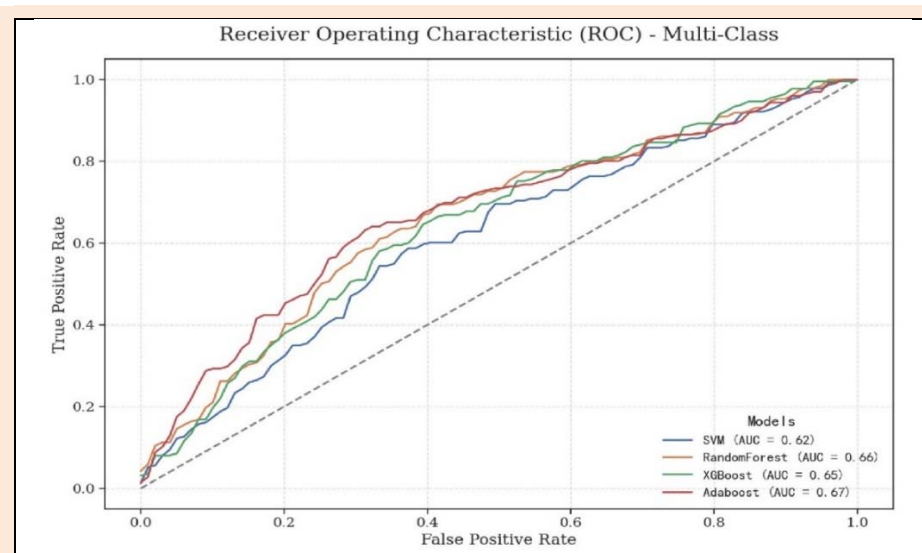


Figure 3. ROC Curves of four algorithms for three-class alertness classification.

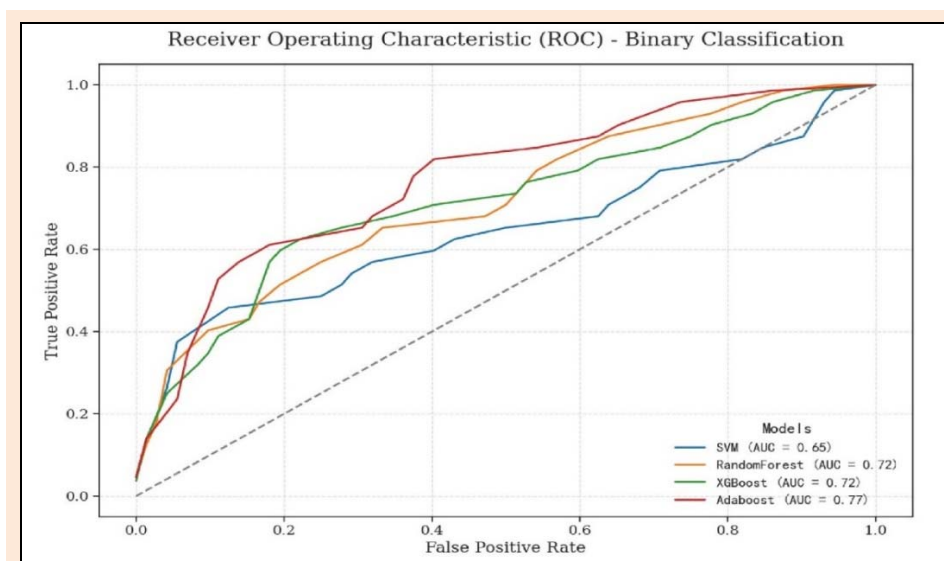
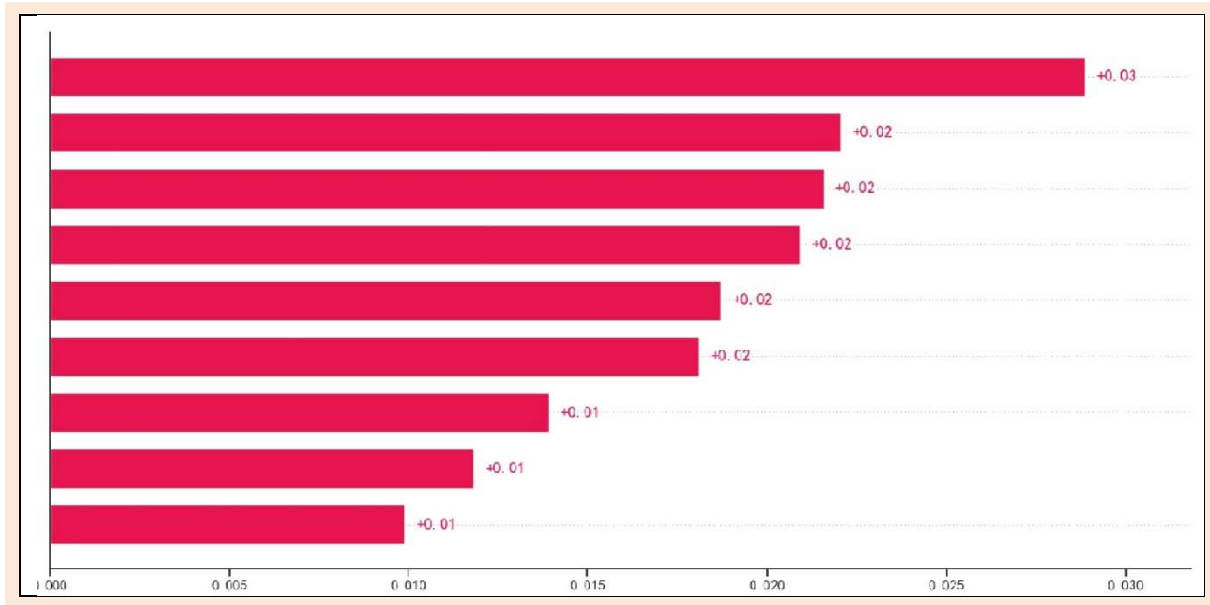
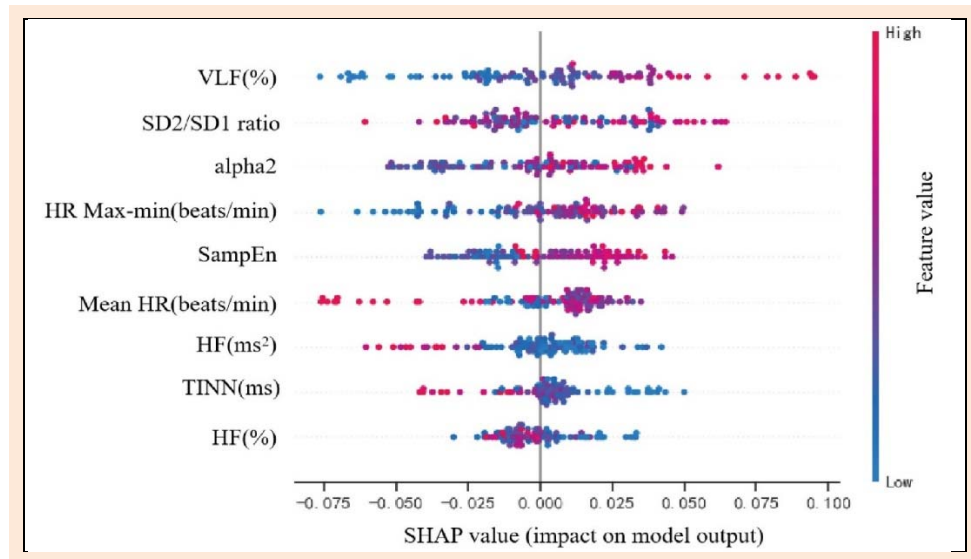


Figure 4. ROC curves of four algorithms for binary alertness classification.



**Figure 5.** SHAP feature importance ranking based on the AdaBoost model.



**Figure 6.** Distribution of each feature's impact on model output.

A key theoretical contribution of this study lies in the identification of specific HRV signatures unique to precision sports. Our SHAP analysis revealed that VLF% (very low frequency percentage) and the SD2/SD1 ratio were the most sensitive predictors of vigilance. The prominence of VLF%, which typically reflects long-term regulatory mechanisms influenced by thermoregulation and hormonal activity, suggests that the physiological demand of shooting differs significantly from high-intensity sports (Storniolo et al., 2025). In the sustained, quasi-isometric state of shooting, slow, integrative physiological rhythms appear crucial for maintaining performance stability. However, the physiological interpretation of VLF% in the context of cognitive alertness warrants careful consideration. Recent methodological guidelines emphasize that HRV indices, despite their accessibility, are notoriously difficult to interpret and can be easily misconstrued (Laborde et al., 2017). The mechanistic basis of VLF is highly complex and heavily

debated; it is significantly influenced by non-cognitive physiological processes, including thermoregulatory, hormonal, and respiratory rhythms, rather than solely reflecting central cognitive vigilance (Quintana et al., 2016). In line with recent calls for nuanced reporting and interpretation in psychophysiological research, while VLF% emerged as a robust statistical predictor in our machine learning framework, it should be viewed as a systemic physiological correlate of the quasi-isometric shooting state rather than a direct, isolated index of cognitive alertness. Similarly, the SD2/SD1 ratio captures the balance between long-term and short-term heart rate dynamics. Its significance indicates that "global autonomic adaptability"—rather than immediate stress reactivity—is the primary determinant of a shooter's capacity to maintain neurovisceral integration under monotonous, high-pressure conditions (Laborde et al., 2017).

While the model's accuracy (0.75) indicates room for refinement, this performance must be interpreted within

the context of the specific cohort. For instance, a recent study employing sliding-window HRV metrics on sleep-deprived healthy adults reported a binary classification accuracy of 89% using SVM (Xie and Ma, 2025). While our AdaBoost model achieved a comparatively lower accuracy of 0.75, this discrepancy highlights the unique physiological profile of elite athletes versus the general population. In sleep deprivation paradigms involving healthy adults, vigilance levels fluctuate drastically, creating distinct, high-amplitude physiological signals that are relatively easy for classifiers to detect. In contrast, elite shooters possess exceptionally fine and stable autonomic regulation, resulting in a 'ceiling effect' where performance and physiology remain consistently high with minimal variance (Laborde et al., 2017; Plews et al., 2017). Consequently, detecting the subtle, micro-level vigilance fluctuations in this highly stable cohort is inherently more challenging. Achieving 75% accuracy under these high-stability conditions therefore demonstrates the robust sensitivity of our proposed framework. Furthermore, compared to traditional frequency-domain-dominated approaches, our feature set—encompassing nonlinear parameters—offers a more comprehensive capture of the complex physiological dynamics in elite athletes. Similar applications in other domains, such as driver sleepiness detection (Persson et al., 2019) and occupational fatigue monitoring, have also reported lower accuracy in trained or habitually alert populations. Regarding demographic generalizability, the present sample was predominantly young (mean age 18.4 years) and mixed-sex without stratification. Given established sex differences in autonomic regulation (Dubol et al., 2021), and the potential influence of training status on HRV profiles, subgroup analyses by sex and experience level represent important directions for future research.

From a practical perspective, this study provides a validated, non-invasive tool for assessing "pre-competition readiness." The use of portable ECG devices improves assessment efficiency by reducing testing time by over 90% compared to behavioral tasks like the PVT (Zhou and Zhang, 2022). These findings align with the trend identified by Reis et al., highlighting the growing role of machine learning in predicting athletic performance (Reis et al., 2024). Although direct application during formal competition is currently constrained by regulations regarding electronic devices (Li et al., 2016), this framework is highly valuable for high-fidelity simulated training. It enables coaches to monitor psychophysiological states in real-time and make data-driven adjustments before athletes enter the competition hall. Future developments may explore minimally obtrusive, regulatory-compliant sensing solutions to bridge the gap between training and competition monitoring (Li et al., 2016).

The present study has several limitations. First, the predictive models were validated only through internal subject-level cross-validation and were not evaluated using an independent external cohort, limiting the generalizability of the findings across populations, sports, and testing conditions. Future studies should prioritize external validation using independent datasets. Second, all

participants were elite shooting athletes ( $n = 83$ ), which, although ensuring high ecological validity, restricts cross-sport generalization. Additionally, no stratified analyses were conducted by sex, age, or training characteristics; such analyses were not feasible given the limited sample size but should be considered in larger cohorts. Third, concurrent subjective alertness measures (e.g., visual analog scale or NASA-TLX) and neurophysiological markers (e.g., EEG) were not included. While this was intended to preserve ecological validity, it precludes assessment of convergent validity and should be addressed in future studies. Fourth, although several strategies were implemented to mitigate overfitting—including subject-level cross-validation, recursive feature elimination, and validation-based hyperparameter tuning—the relatively limited sample size constrains the application of more complex models and the establishment of robust normative references (Collins et al., 2024). In addition, calibration performance of the model was not evaluated in the present study. Calibration curves, which assess the agreement between predicted probabilities and observed outcomes, were not generated. Therefore, the current findings should be interpreted primarily in terms of discrimination performance rather than calibrated probability estimates. Future studies should incorporate calibration analysis (e.g., calibration curves, isotonic regression, or Platt scaling) to enhance the reliability and applicability of model predictions. Furthermore, learning curve analysis was not conducted to evaluate the relationship between sample size and model performance. Although cross-validation results suggested stable model behavior across folds, future work should include learning curve assessment to further verify model robustness and data sufficiency. Fifth, potential confounding factors related to sleep and physiological variability were not fully controlled. Although participants with self-reported sleep disorders were excluded and testing was standardized to the morning, objective assessment of baseline sleep quality (e.g., PSQI or actigraphy) was not conducted, and day-to-day variability in sleep may have influenced HRV and vigilance measures. Similarly, menstrual cycle phase in female athletes was not recorded or controlled, which may have introduced additional variability. Furthermore, while morning testing reduced circadian confounding, diurnal variation in HRV may limit the generalizability of the model to other times of day (Boudreau et al., 2013; Vitale et al., 2019). Finally, post-hoc statistical power analysis was not performed for composite machine learning metrics such as the F1-score, as standardized power frameworks for these indices remain underdeveloped (Collins et al., 2024). Instead, model robustness was inferred from consistent performance across algorithms and validation folds.

## Conclusion

This study identifies the very low frequency percentage (VLF%) and the SD2/SD1 ratio as the most sensitive physiological signatures of vigilance in elite shooting athletes, highlighting the pivotal role of slow-wave autonomic regulation in precision performance. By integrating these features with the AdaBoost algorithm, we

developed a binary classification model that effectively distinguishes optimal from sub-optimal alertness states with superior reliability compared to traditional classifiers. This framework provides a validated, non-invasive, and efficient tool for monitoring pre-competition readiness, offering coaches actionable data to support training optimization in high-fidelity environments.

### Acknowledgements

The study was supported by the Science and Technology Program of the Shanghai Municipal Science and Technology Commission: "Research on New Training Strategies for Improving Athletic Performance in High-Temperature and High-Humidity Environments" (grant number 25Y42800301). The APC was funded by the Shanghai Research Institute of Sports Science (Shanghai Anti-Doping Agency).

The anonymized dataset and analysis code are available upon reasonable written request to the corresponding author, subject to a data sharing agreement compliant with institutional ethics requirements (Ethics approval: LLSC20250005). The authors declare that they have no competing interests.

### References

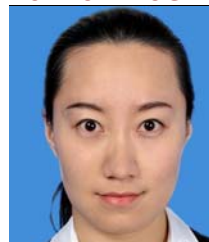
- Balakarhikeyan, V., Jais, R., Vijayarangan, S., Sreelatha Premkumar, P. and Sivaprakasam, M. (2023) Heart Rate Variability Based Estimation of Maximal Oxygen Uptake in Athletes Using Supervised Regression Models. *Sensors (Basel)* **23**. <https://doi.org/10.3390/s23063251>
- Basner, M. and Dinges, D.F. (2011) Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep* **34**, 581-591. <https://doi.org/10.1093/sleep/34.5.581>
- Boudreau, P., Yeh, W.H., Dumont, G.A. and Boivin, D.B. (2013) Circadian variation of heart rate variability across sleep stages. *Sleep* **36**, 1919-1928. <https://doi.org/10.5665/sleep.3230>
- Chang, C.J., Putukian, M., Aerni, G., Diamond, A.B., Hong, E.S., Ingram, Y.M., Reardon, C.L. and Wolanin, A.T. (2020) Mental Health Issues and Psychological Factors in Athletes: Detection, Management, Effect on Performance, and Prevention: American Medical Society for Sports Medicine Position Statement. *Clinical Journal of Sport Medicine* **30**, e61-e87. <https://doi.org/10.1097/jsm.0000000000000817>
- Collins, G.S., Dhiman, P., Ma, J., Schluskel, M.M., Archer, L., Van Calster, B., Harrell, F.E. Jr., Martin, G.P., Moons, K.G.M., van Smeden, M., Sperrin, M., Bullock, G.S. and Riley, R.D. (2024) Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ* **384**, e074819. <https://doi.org/10.1136/bmj-2023-074819>
- Diaz, M.M., Bocanegra, O.L., Teixeira, R.R., Tavares, M., Soares, S.S. and Espindola, F.S. (2013) The relationship between the cortisol awakening response, mood states, and performance. *Journal of Strength and Conditioning Research* **27**, 1340-1348. <https://doi.org/10.1519/jsc.0b013e318267a612>
- Dinges, D.F. and Powell, J.W. (1985) Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers* **17**, 652-655. <https://doi.org/10.3758/bf03200977>
- Dubol, M., Epperson, C.N., Sacher, J., Pletzer, B., Derntl, B., Lanzenberger, R., Sundström-Poromaa, I. and Comasco, E. (2021) Neuroimaging the menstrual cycle: A multimodal systematic review. *Frontiers in Neuroendocrinology* **60**, 100878. <https://doi.org/10.1016/j.yfne.2020.100878>
- Hashemi, M.M., Gladwin, T.E., de Valk, N.M., Zhang, W., Kaldewaij, R., van Ast, V., Koch, S.B.J., Klumpers, F. and Roelofs, K. (2019) Neural Dynamics of Shooting Decisions and the Switch from Freeze to Fight. *Scientific Reports* **9**, 4240. <https://doi.org/10.1038/s41598-019-40917-8>
- Laborde, S., Mosley, E. and Thayer, J.F. (2017) Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research - Recommendations for Experiment Planning, Data Analysis, and Data Reporting. *Frontiers in Psychology* **8**, 213. <https://doi.org/10.3389/fpsyg.2017.00213>
- Li, Q.L., Chen, K.X., Shi, M.Q., Han, Y.L., Chi, L.Z. and Zhou, Y. (2023) Effects of HRV and EEG Biofeedback Training on Athletes with Central Fatigue. *China Sport Science and Technology* **59**, 14-20. <https://doi.org/10.1016/j.ijpsycho.2023.05.244>
- Li, R.T., Kling, S.R., Salata, M.J., Cupp, S.A., Sheehan, J. and Voos, J.E. (2016) Wearable Performance Devices in Sports Medicine. *Sports Health* **8**, 74-78. <https://doi.org/10.1177/1941738115616917>
- Ma, S., Zhang, J., Shi, C., Di, P., Robertson, I.D. and Zhang, Z.Q. (2024) Physics-Informed Deep Learning for Muscle Force Prediction With Unlabeled sEMG Signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **32**, 1246-1256. <https://doi.org/10.1109/tnsre.2024.3375320>
- Mah, C.D., Mah, K.E., Kezirian, E.J. and Dement, W.C. (2011) The effects of sleep extension on the athletic performance of collegiate basketball players. *Sleep* **34**, 943-950. <https://doi.org/10.5665/sleep.1132>
- Malik, M., Bigger, J.T., Camm, A.J., Kleiger, R.E., Malliani, A., Moss, A.J. and Schwartz, P.J. (1996) Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal* **17**, 354-381. <https://doi.org/10.1093/oxfordjournals.eurheartj.a014868>
- Munguía-Izquierdo, D., Segura-Jiménez, V., Camiletti-Moirón, D., Pulido-Martos, M., Alvarez-Gallardo, I.C., Romero, A., Aparicio, V.A., Carbonell-Baeza, A. and Delgado-Fernández, M. (2012) Multidimensional Fatigue Inventory: Spanish adaptation and psychometric properties for fibromyalgia patients. The Al-Andalus study. *Clinical and Experimental Rheumatology* **30**, 94-102. <https://doi.org/10.1186/1471-2474-13-18>
- Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E. and Lindeløv, J.K. (2019) PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* **51**, 195-203. <https://doi.org/10.3758/s13428-018-01193-y>
- Persson, A., Jonasson, H., Fredriksson, I., Wiklund, U. and Ahlstrom, C. (2019) Heart Rate Variability for Driver Sleepiness Classification in Real Road Driving Conditions. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* **2019**, 6537-6540. <https://doi.org/10.1109/embc.2019.8857229>
- Plews, D.J., Scott, B., Altini, M., Wood, M., Kilding, A.E. and Laursen, P.B. (2017) Comparison of Heart-Rate-Variability Recording With Smartphone Photoplethysmography, Polar H7 Chest Strap, and Electrocardiography. *International Journal of Sports Physiology and Performance* **12**, 1324-1328. <https://doi.org/10.1123/ijssp.2016-0668>
- Quintana, D.S., Alvares, G.A. and Heathers, J.A. (2016) Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. *Translational Psychiatry* **6**, e803. <https://doi.org/10.1038/tp.2016.73>
- Reifman, J., Kumar, K., Khitrov, M.Y., Liu, J. and Ramakrishnan, S. (2018) PC-PVT 2.0: An updated platform for psychomotor vigilance task testing, analysis, prediction, and visualization. *Journal of Neuroscience Methods* **304**, 39-45. <https://doi.org/10.1016/j.jneumeth.2018.04.007>
- Reis, F.J.J., Alaiti, R.K., Vallio, C.S. and Hespanhol, L. (2024) Artificial intelligence and Machine Learning approaches in sports: Concepts, applications, challenges, and future perspectives. *Brazilian Journal of Physical Therapy* **28**, 101083. <https://doi.org/10.1016/j.bjpt.2024.101083>
- Shao, M., Lai, Y., Gong, A., Yang, Y., Chen, T. and Jiang, C. (2020) Effect of shooting experience on executive function: differences between experts and novices. *PeerJ* **8**, e9802. <https://doi.org/10.7717/peerj.9802>
- Stornio, J.L., Correale, L., Buzzachera, C.F. and Peyré-Tartaruga, L.A. (2025) Editorial: New perspectives and insights on heart rate variability in exercise and sports. *Frontiers in Sports and Active Living* **7**, 1574087. <https://doi.org/10.3389/fspor.2025.1574087>
- Sun, Z.H., Dai, Y.Y., Jiao, X.J., Jiang, J., Qi, H.Z., Yu, H. and Zhou, P. (2024) EEG-based objective vigilance detection and channel selection techniques. *Manned Spaceflight* **30**, 434-442. <https://doi.org/10.1038/s41598-020-62712-6>
- Tan, C., Wang, J., Cao, G., He, Y., Yin, J., Chu, Y., Geng, Z., Li, L. and Qiu, J. (2023) Psychological changes in athletes infected with Omicron after return to training: fatigue, sleep, and mood. *PeerJ* **11**, e15580. <https://doi.org/10.7717/peerj.15580>
- Torres, C. and Kim, Y. (2019) The effects of caffeine on marksmanship accuracy and reaction time: a systematic review. *Ergonomics* **62**, 1023-1032. <https://doi.org/10.1080/00140139.2019.1613572>

- Van Dongen, H.P., Maislin, G., Mullington, J.M. and Dinges, D.F. (2003) The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep* **26**, 117-126. <https://doi.org/10.1093/sleep/26.2.117>
- Vitale, J.A., Bonato, M., La Torre, A. and Banfi, G. (2019) Heart Rate Variability in Sport Performance: Do Time of Day and Chronotype Play A Role? *Journal of Clinical Medicine* **8**, 723. <https://doi.org/10.3390/jcm8050723>
- Xie, T. and Ma, N. (2025) Tracking vigilance fluctuations in real-time: a sliding-window heart rate variability-based machine-learning approach. *Sleep* **48**, zsae199. <https://doi.org/10.1093/sleep/zsae199>
- Zhou, W.Y. and Zhang, R.L. (2022) Vigilance Level Monitoring Based on Heart Rate Variability and Machine Learning. *Manned Spaceflight* **28**, 779-784. <https://doi.org/10.1093/sleep/zsae199>

### Key points

- Among four machine learning algorithms evaluated, AdaBoost demonstrated superior performance in distinguishing optimal from sub-optimal alertness states via a binary classification framework.
- SHAP analysis identified very low frequency percentage (VLF%) and the SD2/SD1 ratio as the most sensitive HRV predictors of vigilance, highlighting the dominant role of slow-wave autonomic regulation in precision sports performance.
- The proposed framework offers a non-invasive, wearable-compatible tool that reduces psychophysiological assessment time by over 90% compared to behavioral paradigms, providing coaches with actionable data for pre-competition readiness monitoring.
- Findings from 83 national-level athletes indicate that the "ceiling effect" of elite autonomic regulation makes vigilance classification inherently more challenging than in general populations, underscoring the value of sport-specific models for applied performance monitoring.

### AUTHOR BIOGRAPHY



**Jiaojiao LU**  
**Employment**

School of Exercise and Health, Shanghai University of Sport, Shanghai, China.

**Degree**

MSc

**Research interests**

Comprehensive assessment of athletic performance.

**E-mail:** [lujiaojiao1018@163.com](mailto:lujiaojiao1018@163.com)



**Jun QIU**

**Employment**

Shanghai Research Institute of Sports Science (Shanghai Anti-Doping Agency), Shanghai, China.

**Degree**

PhD, Prof.

**Research interests**

Athletic performance monitoring & sports nutrition.

**E-mail:** [qiu jun@shriiss.cn](mailto:qiu jun@shriiss.cn)



**Yan AN**

**Employment**

Shanghai Research Institute of Sports Science (Shanghai Anti-Doping Agency), Shanghai, China.

**Degree**

MSc

**Research interests**

Psychological assessment and mental training for elite athletes.

**E-mail:** [anyan198320@163.com](mailto:anyan198320@163.com)

✉ **Jun Qiu**

Shanghai Research Institute of Sports Science (Shanghai Anti-Doping Agency), Shanghai, 200030, China

### Supplementary Materials

**Supplementary Table 1.** Classification criteria for the five-level alertness framework.

Alertness levels	Assigned value	Number of data points	Percentage (%)	Mean reaction time range (ms)
High alertness	0	120	20.00	< 250.05
Higher alertness	1	120	20.00	250.05~273.983
Moderate alertness	2	120	20.00	273.983~294.658
Lower alertness	3	120	20.00	294.658~319.6
Low alertness	4	120	20.00	≥319.6

**Supplementary Table 2.** Classification criteria for the three-level alertness framework.

Alertness levels	Assigned value	Number of data points	Percentage (%)	Mean reaction time range (ms)
High alertness	0	240	40.00	< 273.983
Moderate alertness	1	120	20.00	273.983~319.6
Low alertness	2	240	40.00	≥319.6

**Supplementary Table 3.** Classification criteria for the binary alertness framework.

Alertness levels	Assigned value	Number of data points	Percentage (%)	Mean reaction time range (ms)
High alertness	0	240	40.00	< 273.983
Low alertness	1	240	40.00	≥319.6