ICON Journal of Engineering Applications of Artificial Intelligence
Volume 02 | Issue 06 | 2026
e-ISSN: 3051-3820
p-ISSN: 3051-3812
Journal homepage: https://iconpublishers.com/icon-j-eng-app-ai/

Research Article

DOI: 10.5281/zenodo.20668114

Machine Learning-Based Shooter Performance Assessment and Intelligent Coaching in Indoor Marksmanship Training Simulators

*** Abubakar Surajo Imam[1], Yakubu Rabiu[2], Mahmud Usman Wubila[3], Aliyu Musa[4], Muhammad Ahmad Baballe[5]**
[1,2,3,4]Department of Mechatronic Engineering, Nigerian Defence Academy, Kaduna, Nigeria.
[5]Department of Mechanical Engineering, Nigerian Defence Academy, Kaduna, Nigeria.

Corresponding author: Abubakar Surajo Imam
Department of Mechatronic Engineering, Nigerian Defence Academy, Kaduna, Nigeria.

Received Date: 02 May 2026
Published Date: 12 June 2026

Abstract
Marksmanship proficiency remains a critical determinant of combat effectiveness in military, law-enforcement, and security operations. Conventional marksmanship training relies heavily on instructor observation and subjective assessment, which may limit consistency, scalability, and objectivity. This paper presents a machine learning-based shooter performance assessment and intelligent coaching framework for indoor marksmanship training simulators. The proposed system integrates computer vision, behavioural analytics, weapon-tracking sensors, and an Intelligent Coaching Recommendation Engine to automatically assess shooter stability, trigger control, aim drift, shot grouping, reaction time, and engagement effectiveness. Multiple learning models, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, Transformer architectures, Random Forests, and XGBoost classifiers, are combined within an ensemble-learning framework for skill classification, performance prediction, and personalised coaching. Experimental validation was conducted using behavioural data collected from 180 participants at the Nigerian Defence Academy Indoor Marksmanship Simulation Centre over a six-month period, comprising 5,200 training sessions and approximately 260,000 recorded shots. Results demonstrated a classification accuracy of 97.1%, precision of 96.5%, recall of 96.2%, F1-score of 96.3%, and area under the curve of 98.1%. The intelligent coaching engine achieved a Coaching Reliability Index of 0.92 and a skill-prediction accuracy of 93.4%, while AI-assisted coaching improved trainee performance by 24.6% compared with conventional instructor-led training. The proposed framework functions as a virtual marksmanship instructor capable of objective performance assessment, automated error diagnosis, adaptive training management, and personalised feedback, thereby enhancing military training effectiveness and scalability.

Keywords: *Marksmanship Training; Computer Vision; Machine Learning; Deep Learning; Human Performance Analytics; Intelligent Coaching; Military Simulation; Behavioural Assessment; Adaptive Training.*

1. Introduction

Marksmanship training remains a fundamental component of military preparedness, despite advances in Artificial Intelligence (AI), autonomous systems, cyber warfare and precision-guided weapons. Personnel involved in conventional, asymmetric, counter-terrorism, urban and peace-support operations still require accurate and timely target-engagement skills for mission success and survivability [1]–[4]. Traditional marksmanship instruction depends largely on experienced instructors who assess weapon handling, trigger discipline, aiming stability, recoil management, shot grouping, reaction time and engagement effectiveness during live-fire or simulated exercises [5], [6]. Although effective, this approach is often subjective and influenced by instructor experience, observation quality, fatigue and trainee-to-instructor ratio. As training populations increase, maintaining consistent and objective assessment becomes more difficult.

Modern indoor marksmanship simulators now incorporate high-speed cameras, laser tracking, motion sensors, computer-generated scenarios, and automated scoring systems [7]–[10]. These technologies generate large volumes of behavioural data and create opportunities to move beyond score-based evaluation towards intelligent performance assessment. AI and Machine Learning (ML) provide powerful tools for analysing such data, particularly in computer vision, behavioural recognition, activity classification and predictive analytics [11]–[20]. The proposed framework applies AI as a virtual marksmanship instructor capable of continuously monitoring shooter behaviour, diagnosing deficiencies, classifying skill level, predicting progression, and recommending personalised corrective actions. Behavioural indicators such as weapon stability, trigger manipulation, aim drift, recoil recovery, reaction time, and shot grouping are processed through machine-learning models to support adaptive coaching [21]–[25]. Figure 1 presents the conceptual architecture of the intelligent coaching framework.

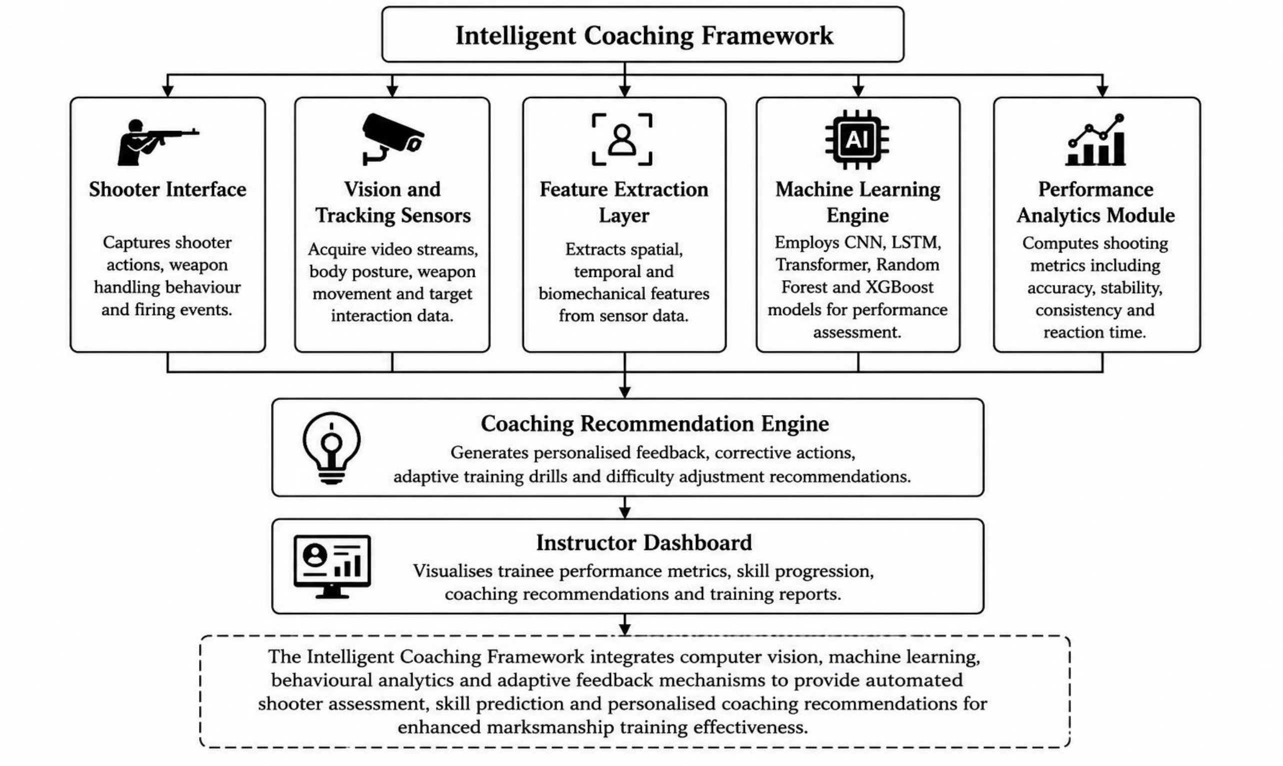


Fig. 1: Conceptual Architecture of the Intelligent Coaching Framework

AI-enabled coaching offers advantages over conventional instructor-centric assessment, particularly in scalability, objectivity, continuous monitoring, personalised feedback, performance prediction and adaptive difficulty control. These benefits are especially relevant for military organisations required to train large numbers of personnel while maintaining standardised performance assessment.

Table 1: Comparison Between Conventional and AI-Enabled Coaching Systems

Attribute	Conventional Assessment	AI-Based Assessment
Consistency	Instructor dependent	High
Scalability	Limited	High
Continuous Monitoring	Limited	Yes
Personalised Feedback	Moderate	High
Performance Prediction	No	Yes
Adaptive Difficulty	No	Yes
Data Analytics	Limited	Extensive

The computational foundation of the proposed system is summarised in Table 2.

Table 2: Machine-Learning Techniques Used in Intelligent Coaching

Model	Primary Function
CNN	Spatial feature extraction
LSTM	Temporal behaviour modelling
Transformer	Long-range sequence analysis
Random Forest	Behaviour classification
XGBoost	Predictive analytics
Ensemble Learning	Decision fusion

Although previous studies have examined automated scoring, shooter tracking, target detection, and machine-learning-assisted training, limited work has developed a unified intelligent coaching framework that simultaneously evaluates shooter stability, trigger control, aim drift, shot grouping, skill level, and training progression [26]–[35]. This paper therefore proposes a ML-based shooter performance assessment and intelligent coaching framework for indoor marksmanship training simulators. The principal contributions of this work are as follows:

- Development of a computer vision-based framework for automated shooter-behaviour assessment.
- Quantitative evaluation of weapon stability, trigger control, aim drift, and shot-grouping characteristics.
- Integration of CNN, LSTM, Transformer, Random Forest, and XGBoost models within a unified ensemble-learning architecture.
- Development of an Intelligent Coaching Recommendation Engine capable of automated error diagnosis, root-cause analysis, and personalised corrective feedback.
- Implementation of adaptive training and skill-progression prediction mechanisms.
- Experimental validation using large-scale behavioural data collected at the Nigerian Defence Academy Indoor Marksmanship Simulation Centre.

The remainder of the paper is organised as follows. Section 2 reviews related work. Section 3 presents the system architecture and data acquisition framework. Sections 4–7 describe computer-vision processing, machine-learning models, and coaching methodologies. Sections 8 and 9 present results, operational implications, robustness, limitations, and conclusions.

2. Related Work and Literature Review

Marksmanship training remains a critical component of military readiness, particularly in conventional warfare, counter-terrorism, force protection, and close-quarter combat operations. Traditional assessment methods focus primarily on observable outcomes such as hit probability, shot placement, shot grouping, reaction time, and instructor judgement. While these metrics provide useful indicators of shooting effectiveness, they offer limited insight into the behavioural factors responsible for performance outcomes. Recent advances in computer vision, sensor fusion, ML and human-performance analytics have shifted assessment from outcome-based scoring to behaviour-based evaluation. Modern simulators can analyse weapon stability, trigger control, aim drift, recoil recovery, and target-engagement behaviour, thereby enabling more objective and comprehensive performance assessment. Although commercial systems provide automated scoring and after-action review capabilities, few offer detailed behavioural diagnosis and intelligent coaching. Consequently, opportunities remain for developing AI-enabled systems capable of functioning as virtual marksmanship instructors [36]–[42].

2.1 Computer Vision for Shooter Behaviour Analysis

Computer vision has emerged as a key enabling technology for automated shooter-performance assessment. It enables cameras and image-processing algorithms to monitor shooter posture, weapon movement, aiming behaviour, recoil recovery, and shot-impact locations. These capabilities allow evaluation not only of where shots land but also of how the shooter behaves before and after firing. Recent advances in deep learning, object detection, human pose estimation, and action recognition have significantly improved the capability of computer-vision systems to analyse human behaviour in real time [43]–[48]. Fig. 2 illustrates the evolution of computer-vision applications in marksmanship training.

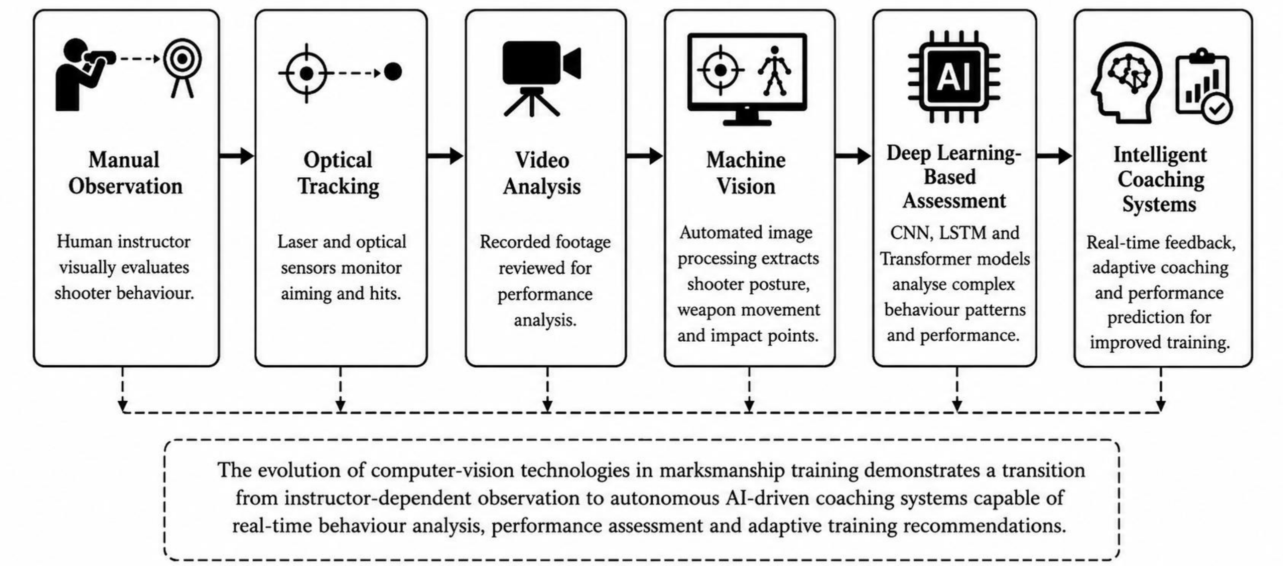


Fig. 2: Evolution of Computer Vision Applications in Marksmanship Training.

Deep-learning techniques, particularly Convolutional Neural Networks (CNNs), have significantly improved behavioural analysis by automatically learning spatial features from image sequences. CNNs are especially suitable for posture analysis, weapon-alignment monitoring, stability assessment, and movement tracking [49]–[53].

$$y(i,j) = \sum_{m}\sum_{n} x(i-m, j-n)k(m,n) \qquad (1)$$

where x is the input image, $k$is the convolution kernel, and y is the resulting feature map.

2.2 Human Performance Analytics in Military Training

Human-performance analytics focuses on measuring and interpreting behavioural, cognitive, and physical factors that influence operational effectiveness. In marksmanship training, these factors include reaction time, situational awareness, cognitive workload, trigger discipline, aim stability, and recoil recovery. Frameworks such as Endsley's Situational Awareness Model and the NASA Task Load Index remain widely adopted for evaluating decision-making and workload during military training activities [54]–[58]. Recent military human-performance studies have increasingly incorporated physiological sensing, eye tracking, wearable sensors, and cognitive-state monitoring to improve assessment accuracy and training effectiveness [59]–[62].

Table 3: Human Performance Indicators Relevant to Marksmanship Training

Indicator	Description
Accuracy	Target-hit effectiveness
Precision	Shot consistency
Reaction Time	Engagement speed
Situational Awareness	Environmental understanding
Cognitive Workload	Mental effort
Recoil Recovery	Post-shot stabilisation
Trigger Discipline	Trigger-control quality
Aim Stability	Weapon steadiness

As shown in Table 3, effective marksmanship assessment requires a multidimensional evaluation framework extending beyond simple shot scoring.

2.3 Machine Learning for Behavioural Assessment

Machine learning enables automated identification of behavioural patterns within complex shooter-performance datasets and has become increasingly important for shooter classification, error detection, performance prediction, and personalised coaching. Supervised-learning algorithms such as Random Forest, Support Vector Machine, Gradient Boosting, and XGBoost have demonstrated strong performance in human-activity recognition, sports analytics, and military-training applications [63]–[67]. Random Forest and XGBoost are particularly effective because they provide robust classification performance and relatively interpretable outputs.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t\,(x) \qquad (2)$$

where $h_t(x)$ represents the prediction of the t^{th} decision tree and $T$is the number of trees. These methods are suitable for classifying shooters into proficiency levels and identifying behavioural deficiencies requiring corrective coaching [68]–[70].

2.4 Deep Learning for Temporal Behaviour Modelling

Many shooting behaviours evolve over time and therefore require temporal modelling. Examples include aim drift, trigger squeeze, recoil recovery, and target-transition behaviour. Long Short-Term Memory (LSTM) networks and Transformer architectures have demonstrated strong performance in modelling such sequential patterns [71]–[75].

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \qquad (3)$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (4)$$

These architectures support behavioural prediction, aim-drift detection, trigger-control analysis, recoil-recovery modelling, and skill-progression assessment. Recent studies have shown that Transformer-based architectures often outperform conventional recurrent networks for long-duration behavioural analysis tasks [76]–[79].

2.5 Intelligent Tutoring and Adaptive Coaching Systems

Intelligent Tutoring Systems (ITS) provide personalised instruction by continuously assessing learner performance and adapting training content accordingly. Within military training environments, such systems can reduce instructor workload while improving feedback consistency and training effectiveness [80]–[83]. Recent advances in explainable AI, reinforcement learning, adaptive learning systems, and human-centred AI have accelerated the development of intelligent coaching systems capable of generating personalised recommendations and adaptive training pathways [84]–[87].

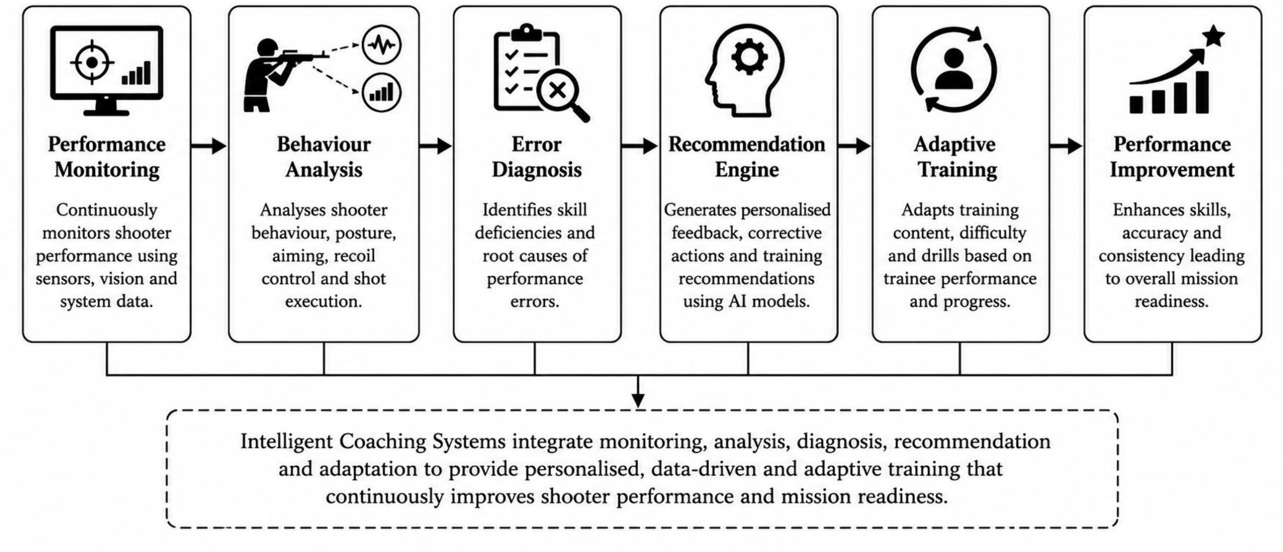


Fig. 3: General Architecture of Intelligent Coaching Systems

The framework creates a closed-loop training process in which behavioural assessment directly informs corrective actions and adaptive training adjustments. Benefits include objective assessment, personalised coaching, real-time feedback, and improved learning efficiency.

2.6 Research Gap Analysis

Despite significant advances in computer vision, ML and intelligent tutoring systems, relatively limited research has focused on fully integrated AI-based coaching systems for military marksmanship training. Existing studies generally address isolated capabilities such as shot scoring, target engagement, shooter tracking, or posture analysis rather than comprehensive behavioural assessment and adaptive coaching [88]–[92].

Table 4: Research Gap Analysis

Capability	Existing Studies	Gap
Shot Scoring	Extensive	Mature
Shooter Stability Analysis	Limited	Significant
Trigger-Control Assessment	Limited	Significant
Aim Drift Detection	Moderate	Moderate
Skill Prediction	Limited	Significant
Intelligent Coaching	Limited	Significant
Adaptive Training	Limited	Significant
Multi-Model AI Fusion	Rare	Major Gap

As shown in Table 4, significant gaps remain in intelligent coaching, adaptive training and integrated behavioural assessment. The present study addresses these limitations through a unified machine-learning framework that combines computer vision, behavioural analytics, multi-model AI fusion, skill classification, performance prediction, and adaptive coaching within a single intelligent marksmanship-training architecture.

3. System Architecture and Data Acquisition Framework

The proposed ML-Based Shooter Performance Assessment and Intelligent Coaching System was developed as a virtual marksmanship instructor capable of continuously monitoring shooter behaviour, diagnosing performance deficiencies, predicting skill progression, and generating personalised coaching recommendations. Unlike conventional simulator systems that focus primarily on hit detection and score computation, the framework integrates multi-modal sensing, computer vision, behavioural analytics, machine learning, and adaptive coaching within a unified architecture. Such integrated AI-enabled training systems have recently attracted considerable interest because they provide objective performance assessment, reduce instructor workload, and enable personalised learning pathways [93]–[97].

By analysing both shooting outcomes and behavioural processes, the proposed framework provides objective, repeatable, and data-driven assessment while reducing dependence on human instructors. Similar intelligent assessment architectures have demonstrated effectiveness in military simulation, sports-performance analysis, and human-performance optimisation applications [98]–[101]. The overall architecture of the framework is presented in Fig. 4. Behavioural data acquired from multiple sensing systems are fused and analysed by machine-learning algorithms capable of identifying shooter errors and recommending corrective actions in real time. The architecture is intentionally modular to support future integration of physiological monitoring systems, virtual reality (VR), augmented reality (AR), mixed reality (MR), and digital-twin training environments [102]–[105].

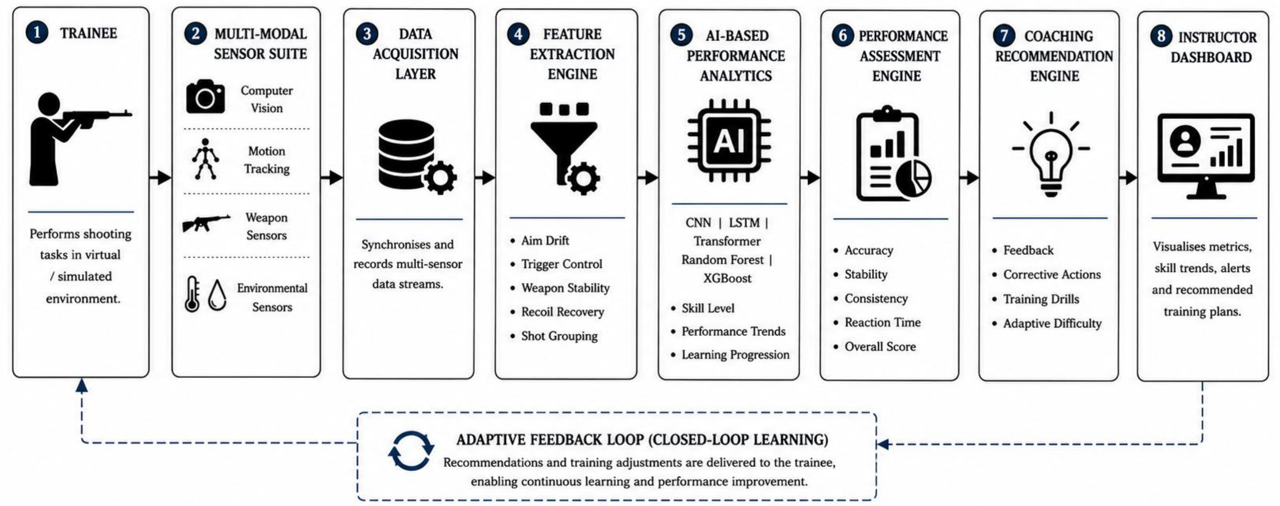


Fig. 4: Overall Architecture of the Intelligent Coaching Framework

The framework consists of six principal subsystems whose functions are summarised in Table 5.

Table 5: Functional Components of the Proposed System

Module	Function
Data Acquisition	Collection of behavioural and performance data
Computer Vision	Posture, movement and weapon tracking
Feature Extraction	Generation of machine-learning features
ML Analytics	Classification, prediction and assessment
Coaching Engine	Generation of corrective recommendations
Instructor Dashboard	Visualisation and reporting

As shown in Table 5, each subsystem transforms raw sensor observations into actionable performance indicators that support intelligent coaching and adaptive training.

3.1 Data Acquisition Framework

Reliable behavioural assessment requires comprehensive acquisition of shooter-performance information. Traditional marksmanship systems typically record only target-hit locations and shot timing, providing limited insight into the behavioural processes that influence performance. To address this limitation, the proposed framework employs a multi-sensor architecture that continuously monitors weapon motion, aiming behaviour, trigger manipulation, and engagement dynamics throughout the firing sequence. Recent advances in sensor fusion have demonstrated that combining multiple sensing modalities improves robustness, measurement accuracy, and behavioural understanding compared with single-sensor systems [106]–[109]. Accordingly, the proposed framework integrates optical, inertial, mechanical, and ballistic sensing technologies within a unified acquisition architecture. The overall data-acquisition architecture is illustrated in Fig. 5.

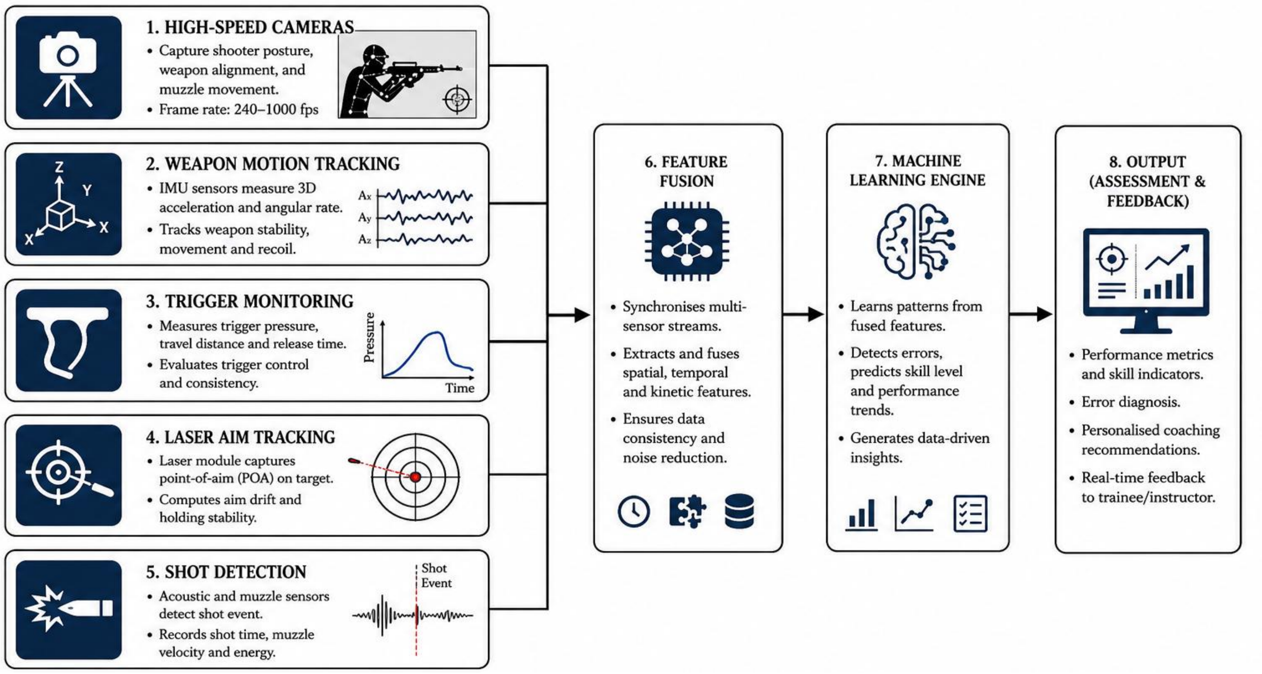


Fig. 5: Multi-Sensor Data Acquisition Framework

The principal categories of acquired behavioural information are summarised in Table 6.

Table 6: Acquired Shooter Behaviour Data

Data Category	Measured Variable
Aim Behaviour	Aim-point coordinates
Weapon Motion	Angular displacement
Trigger Behaviour	Pull characteristics
Shot Outcomes	Impact coordinates
Engagement Timing	Reaction time
Target Interaction	Engagement effectiveness
Historical Records	Previous performance

The complete behavioural feature vector is represented as:

$$\mathbf{x} = [x_1, x_2, \ldots, x_n] \qquad (5)$$

where x_i denotes an individual behavioural feature extracted from sensor observations. The feature vector contains information relating to shooter stability, trigger control, aim drift, shot grouping, reaction time, recoil recovery, and engagement effectiveness.

3.2 Computer Vision Subsystem

Computer vision constitutes one of the most important components of the proposed framework because many critical shooter behaviours are visually observable. Recent advances in deep learning, object detection, human pose estimation, and motion tracking have significantly improved the capability of computer-vision systems to analyse complex human activities in real time [110]–[114].

The subsystem employs four synchronised high-speed cameras operating at 240 fps to monitor shooter posture, weapon alignment, aiming behaviour, trigger manipulation, and recoil recovery. Multi-camera configurations have been shown to improve tracking accuracy and reduce occlusion effects in behavioural-analysis applications [115], [116]. The camera deployment configuration is illustrated in Fig. 6.

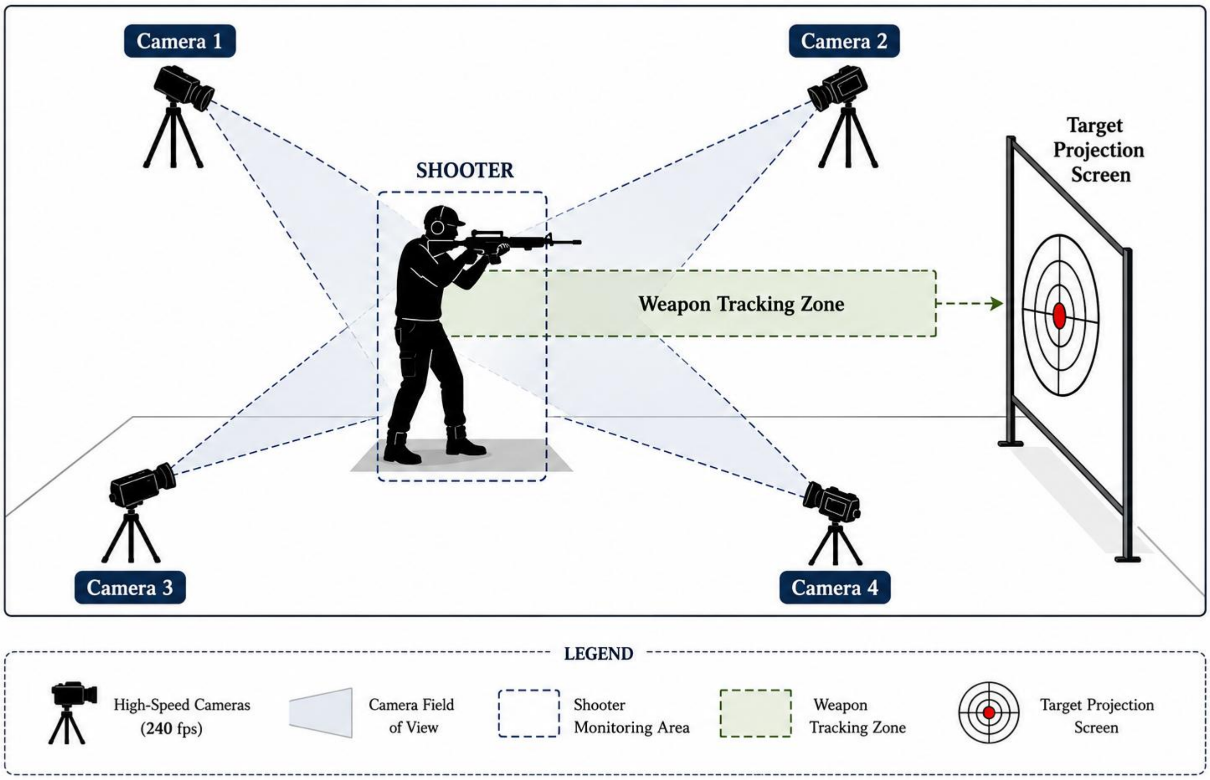


Fig. 6: Camera Deployment for Shooter Monitoring.

The multi-camera arrangement minimises occlusion and improves measurement accuracy by providing multiple viewing perspectives. Camera calibration is performed using the projective camera model:

$$s\begin{bmatrix}u\\v\\1\end{bmatrix} = K[R \mid t]\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix} \qquad (6)$$

where $K$is the intrinsic camera matrix, $R$is the rotation matrix, $t$is the translation vector, and $s$is the scale factor. Similar calibration models are widely used in multi-view computer-vision systems for three-dimensional scene reconstruction and motion tracking [117], [118].

3.3 Feature Extraction and Behaviour Representation

Behavioural observations acquired from the computer-vision subsystem, weapon-tracking sensors, trigger sensors, and simulator software are transformed into machine-learning-compatible representations through feature extraction. Feature engineering is a critical stage because classification accuracy depends heavily on the quality and relevance of the extracted features [119]–[121]. The overall feature-extraction process is illustrated in Fig. 7. As shown in the figure, multimodal data from video streams, weapon tracking, trigger sensors, and shot records are processed to generate behavioural indicators describing shooter performance. These indicators are subsequently normalised and supplied to the machine-learning models.

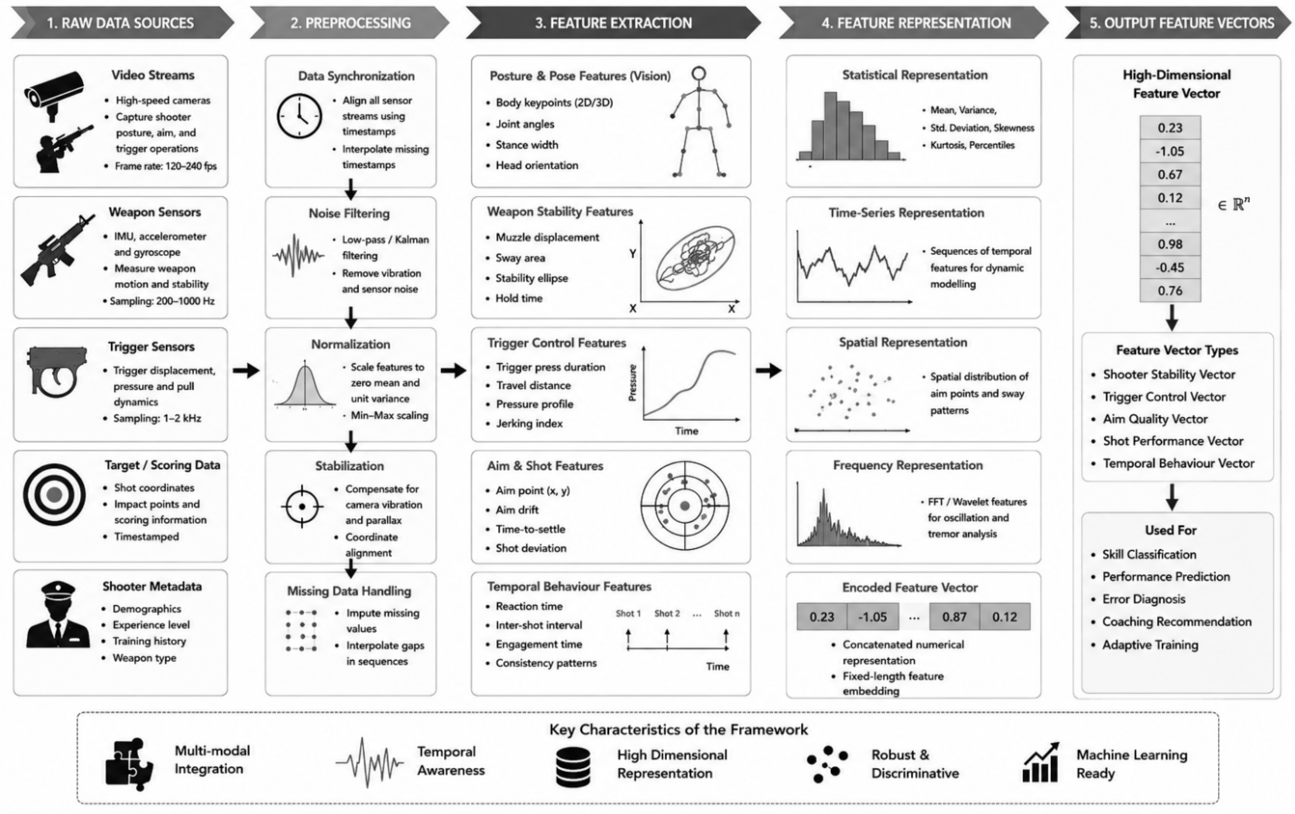


Fig. 7: Behavioural Feature Extraction and Representation Framework

The behavioural features extracted by the framework are summarised in Table 7.

Table 7: Behavioural Features Extracted for Machine Learning

Feature	Description
Stability Index	Weapon steadiness
Trigger Smoothness	Trigger-control quality
Aim Drift	Mean aiming deviation
Shot Group Radius	Shot consistency
Reaction Time	Engagement speed
Recovery Time	Recoil recovery behaviour
Engagement Accuracy	Target-hit effectiveness
Decision Quality	Correct target selection

As shown in Table 7, the extracted features capture the key behavioural characteristics associated with marksmanship performance. Prior to machine-learning analysis, features are normalised using z-score normalisation:

$$z_i = \frac{x_i - \mu_i}{\sigma_i} \qquad (7)$$

where x_i is the raw feature value, μ_i is the feature mean, and $\sigma_i$is the feature standard deviation. Normalisation improves training stability and ensures that individual features contribute appropriately during model optimisation [122]. The resulting feature vector forms the primary input to the machine-learning framework described in the subsequent sections.

3.4 Dataset Construction

A comprehensive dataset was constructed from indoor marksmanship training sessions involving both novice (cadets) and experienced shooters. The dataset includes video streams, weapon-motion data, trigger signals, aim trajectories, shot coordinates, and performance records. The use of diverse datasets containing multiple proficiency levels improves model generalisation and reduces bias during classifier training [123], [124]. The characteristics of the acquired dataset are summarised in Table 8. As shown in Table 8, the dataset comprises 180 participants, 5,200 training sessions, and approximately 260,000 recorded shots, providing a substantial volume of behavioural information for machine-learning model development. The dataset contains 32 behavioural features extracted from four synchronised camera streams operating at 240 fps and represents five distinct shooter skill classes.

Table 8: Dataset Characteristics

Parameter	Value
Participants	180
Training Sessions	5,200
Total Shots	260,000
Behavioural Features	32
Skill Classes	5
Camera Streams	4
Sampling Rate	240 fps

To ensure unbiased model evaluation and minimise overfitting, the dataset was partitioned into training, validation and testing subsets as presented in Table 9. Seventy per cent of the dataset was allocated for model training, while 15% was used for hyperparameter optimisation and validation, and the remaining 15% was reserved for independent performance testing.

Table 9: Dataset Partitioning

Dataset	**Percentage**
Training	70%
Validation	15%
Testing	15%

The dataset size, diversity and partitioning strategy shown in Tables 8 and 9 provide a robust foundation for machine-learning training, model optimisation, and statistically meaningful performance evaluation.

3.5 Performance Assessment Workflow

The operational workflow of the intelligent coaching framework is illustrated in Fig. 8.

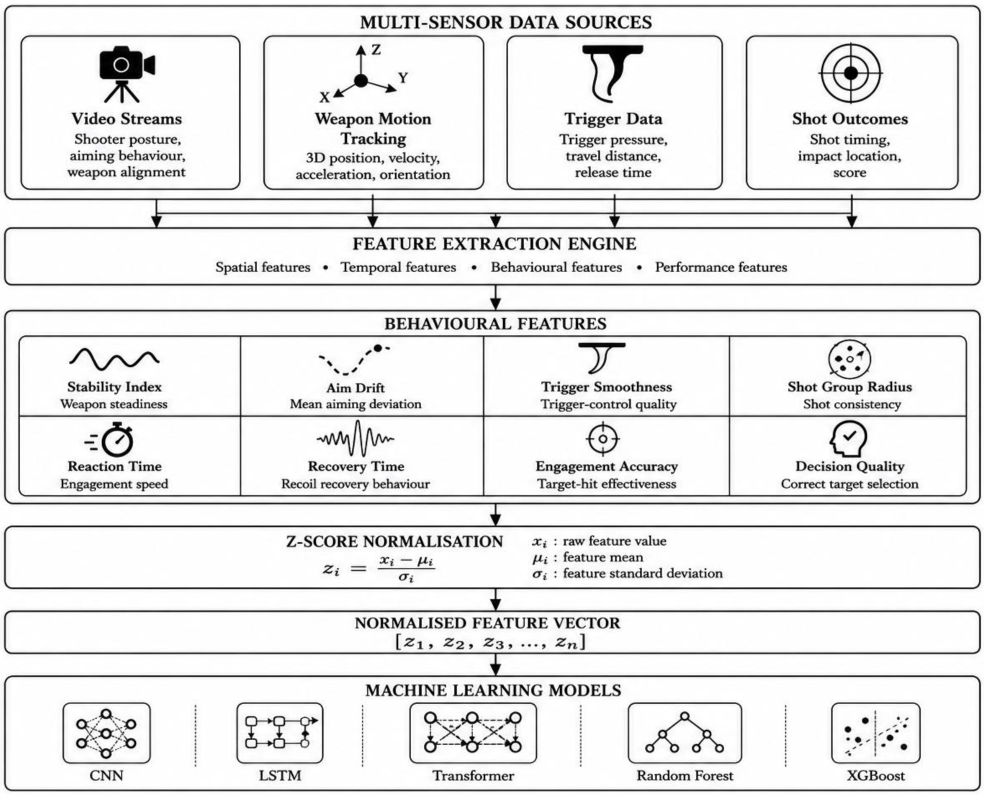


Fig. 8: Behavioural Assessment Workflow

The workflow shown in Fig. 8 forms the operational backbone of the proposed system. Data acquired from multiple sensors are transformed into behavioural features, analysed using machine-learning models, and converted into personalised coaching recommendations. Similar closed-loop assessment architectures have demonstrated substantial improvements in training efficiency, adaptive learning, and performance optimisation across defence, sports, and aviation-training applications [125]–[128].

Through continuous assessment and adaptive feedback, the proposed framework supports accelerated skill acquisition, objective performance evaluation, and improved marksmanship proficiency.

4. Computer Vision-Based Shooter Behaviour Analysis

Computer vision enables objective assessment of shooter performance by transforming video streams and sensor measurements into quantitative behavioural indicators. Unlike conventional marksmanship evaluation, which primarily measures target-hit outcomes, the proposed framework analyses weapon stability, trigger control, aim drift, and shot consistency to identify performance deficiencies and support intelligent coaching. Recent advances in deep learning, object tracking, pose estimation, and motion analysis have significantly improved the capability of computer-vision systems to extract behavioural information in real time [129]–[135].

The behavioural-analysis framework is presented in Fig. 9. As shown in Fig. 9, video streams and sensor data are processed through object-tracking and feature-extraction modules to generate performance indicators for machine-learning-based assessment. The extracted indicators are summarised in Table 10.

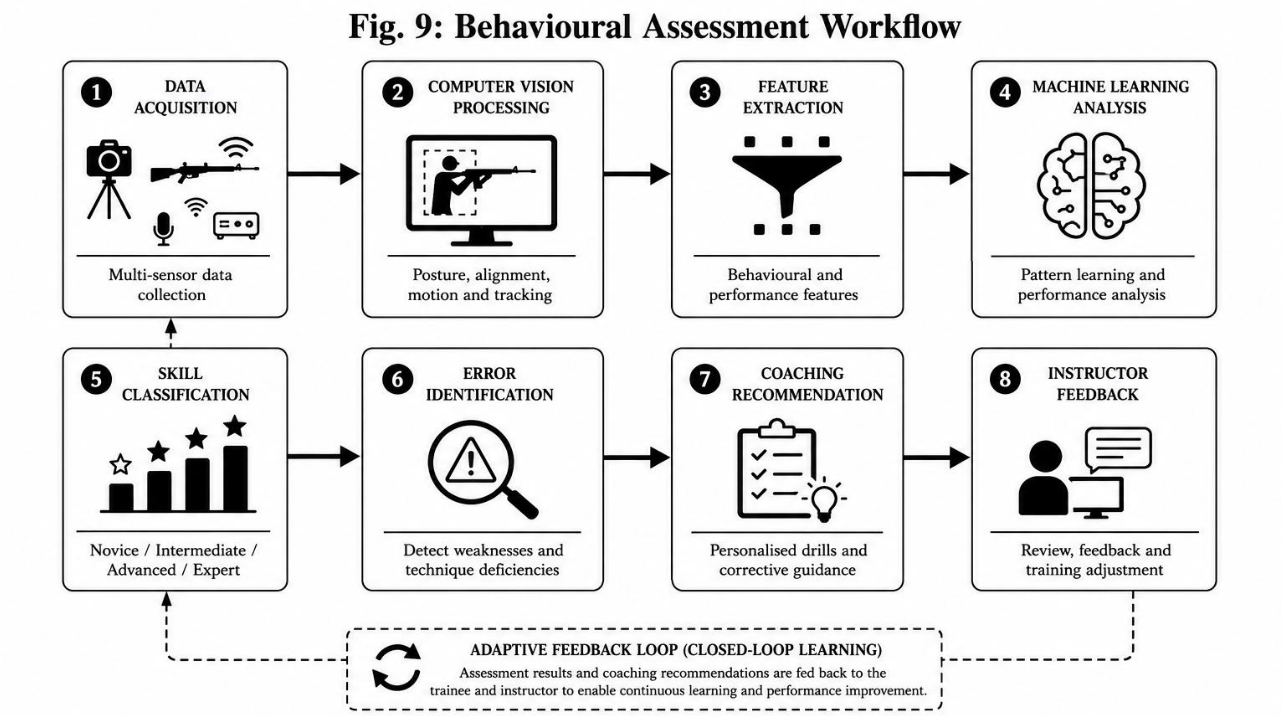


Fig. 9: Computer Vision-Based Behaviour Analysis Framework

Table 10: Behavioural Indicators Used for Performance Assessment

Indicator	Description
Stability Index	Weapon steadiness
Trigger Quality	Trigger-control effectiveness
Aim Drift	Aim-point deviation
Shot Group Radius	Fire consistency
Recovery Time	Recoil stabilisation
Reaction Time	Engagement speed
Engagement Accuracy	Target-hit effectiveness

As shown in Table 10, these indicators collectively capture the key behavioural characteristics of marksmanship performance and provide the basis for automated assessment and coaching.

4.1 Shooter Stability Assessment

Weapon stability significantly influences shooting accuracy because excessive weapon movement increases shot dispersion. The proposed framework tracks aim-point trajectories and evaluates stability using the standard deviation of aiming displacement. The assessment procedure is illustrated in Fig. 10, while the classification criteria are summarised in Table 11.

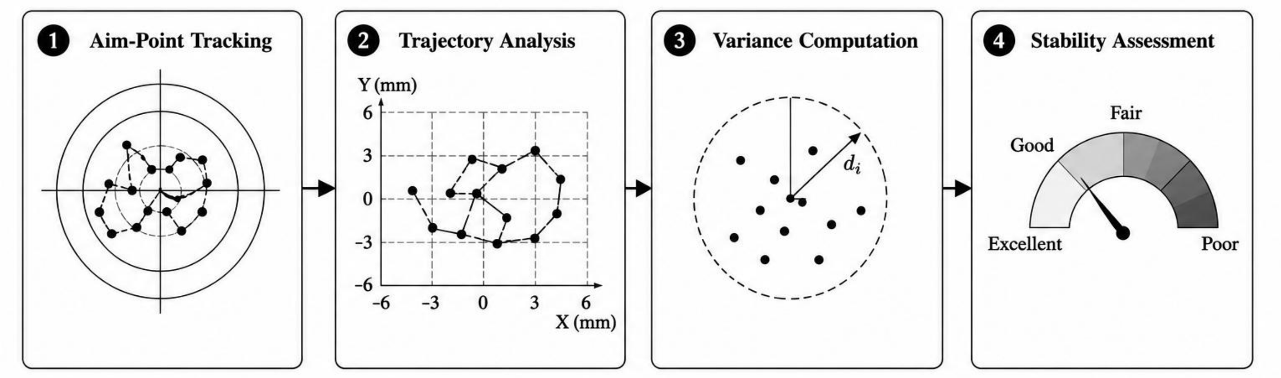


Fig. 10: Weapon Stability Assessment

Aim-point variability is quantified as:

$$\sigma_d = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - \bar{d})^2} \qquad (8)$$

where: d_i= instantaneous aiming displacement, $\bar{d}$= mean aiming displacement and N= number of samples.

Table 11: Stability Assessment Criteria

Standard Deviation	Assessment
< 1.0 mm	Excellent
1.0–2.5 mm	Good
2.5–5.0 mm	Fair
> 5.0 mm	Poor

As shown in Table 11, increasing aim-point dispersion corresponds to poorer weapon control and reduced shooting effectiveness.

4.2 Trigger Control Evaluation

Trigger manipulation directly affects shot placement because abrupt trigger activation can disturb weapon alignment. The trigger-assessment framework is illustrated in Fig. 11, while the evaluation criteria are presented in Table 12.

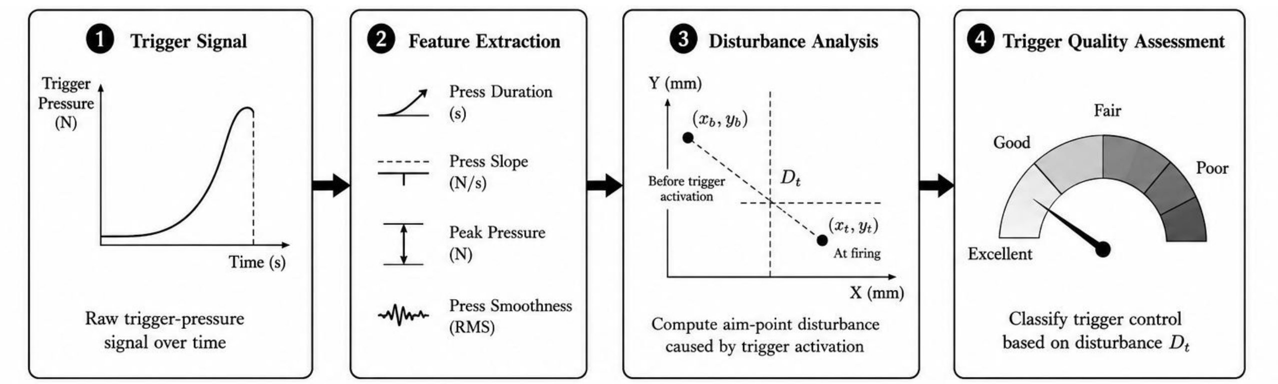


Fig. 11: Trigger-Control Assessment Framework

Trigger disturbance is computed as:

$$D_t = \sqrt{(x_t - x_b)^2 + (y_t - y_b)^2} \qquad (9)$$

where: (x_b, y_b)= aim coordinates immediately before trigger activation and (x_t, y_t)= aim coordinates at firing.

Table 12: Trigger-Control Assessment Criteria

Trigger Disturbance	Assessment
< 1.0 mm	Excellent
1.0–2.5 mm	Good
2.5–5.0 mm	Fair
> 5.0 mm	Poor

Lower values of D_t indicate superior trigger discipline and reduced weapon disturbance during firing.

4.3 Aim Drift Detection

Aim drift refers to gradual displacement of the aiming point caused by fatigue, breathing irregularities, poor posture, or psychological stress. Continuous monitoring provides an objective measure of aiming consistency and concentration. The mean drift magnitude is computed as:

$$AD = \frac{1}{N}\sum_{i=1}^{N} \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \qquad (10)$$

where:
(x_0, y_0)= desired aiming location,
(x_i, y_i)= instantaneous aim-point coordinates.
The maximum drift magnitude is defined as:

$$AD_{\max} = \max\left(\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}\right) \qquad (11)$$

The adopted drift-assessment criteria are summarised in Table 13.

Table 13: Aim Drift Assessment Criteria

Drift Magnitude	Assessment
< 2 mm	Excellent
2–5 mm	Good
5–10 mm	Fair
> 10 mm	Poor

As shown in Table 13, increasing drift generally corresponds to declining aiming consistency and reduced marksmanship performance.

4.4 Shot Grouping Analysis

Shot grouping provides a direct measure of shooting consistency and repeatability. While engagement accuracy indicates target-hit effectiveness, shot-group characteristics reveal the quality of weapon control maintained throughout the firing sequence. The analysis process is illustrated in Fig. 12.

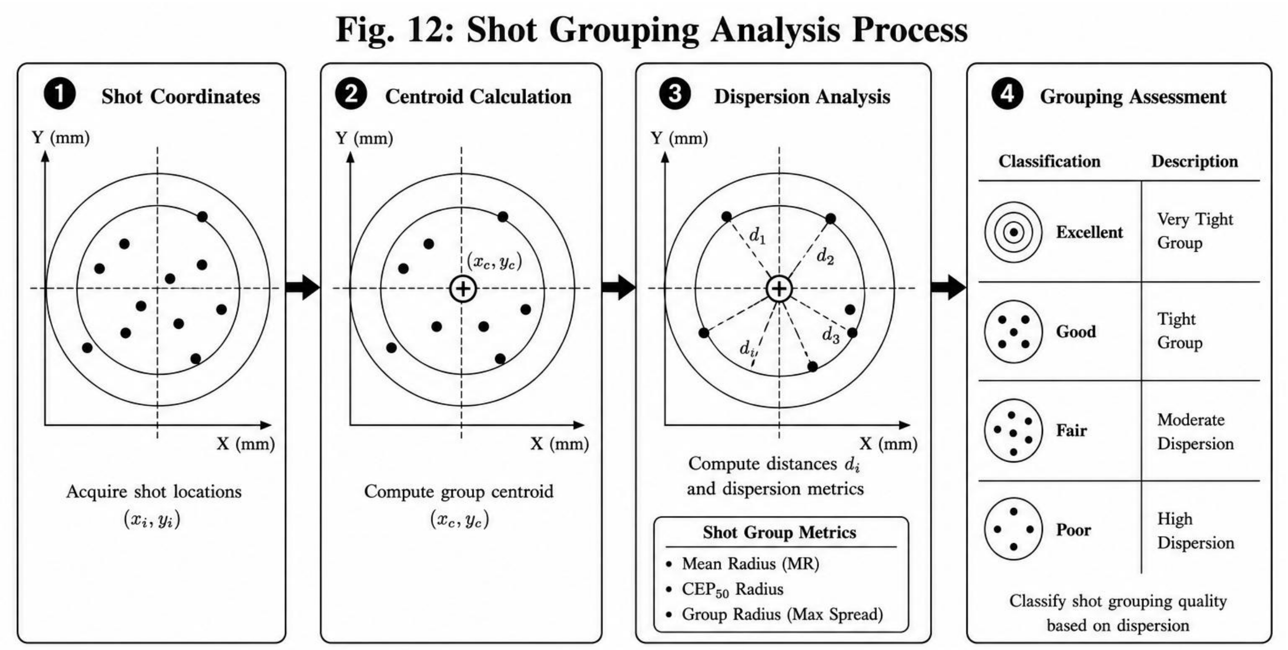


Fig. 12: Shot Grouping Analysis Process

The shot-group centroid is determined as:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (12)$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad (13)$$

The shot-group radius is then computed as:

$$SG = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(x_i - \bar{x})^2 + (y_i - \bar{y})^2]} \qquad (14)$$

where: SG= shot-group radius and $(\bar{x}, \bar{y})$= shot-group centroid. Smaller values of SGindicate superior shooting consistency and weapon control.

4.5 Behavioural Feature Fusion

The behavioural indicators generated from stability assessment, trigger evaluation, aim-drift detection, and shot-group analysis are fused into a unified feature representation for machine-learning analysis:

$$\mathbf{F} = [\sigma_d, \; D_t, \; AD, \; SG, \; RT, \; RR, \; EA] \qquad (15)$$

where: σ_d= stability measure, D_t= trigger disturbance, AD= mean aim drift, SG= shot-group radius, RT= reaction time, RR= recoil-recovery time and EA= engagement accuracy.

The fused feature vector forms the primary input to the machine-learning framework presented in the next section. By integrating multiple behavioural indicators, the proposed system provides a comprehensive representation of shooter performance suitable for automated assessment, skill classification, performance prediction, and intelligent coaching.

5. Machine Learning Framework for Shooter Classification and Intelligent Coaching

The behavioural indicators developed in Section 4 provide quantitative representations of shooter performance. However, transforming these measurements into actionable coaching recommendations requires ML models capable of identifying complex relationships between behavioural characteristics and marksmanship proficiency. Recent advances in deep learning, ensemble learning, and attention-based architectures have demonstrated strong performance in behavioural analytics, human activity recognition, and intelligent training systems [136]–[142]. The proposed ML framework is illustrated in Fig. 13. As shown in Fig. 13, behavioural features extracted from the computer-vision subsystem are processed through complementary learning models to perform skill classification, performance prediction, and coaching recommendation generation.

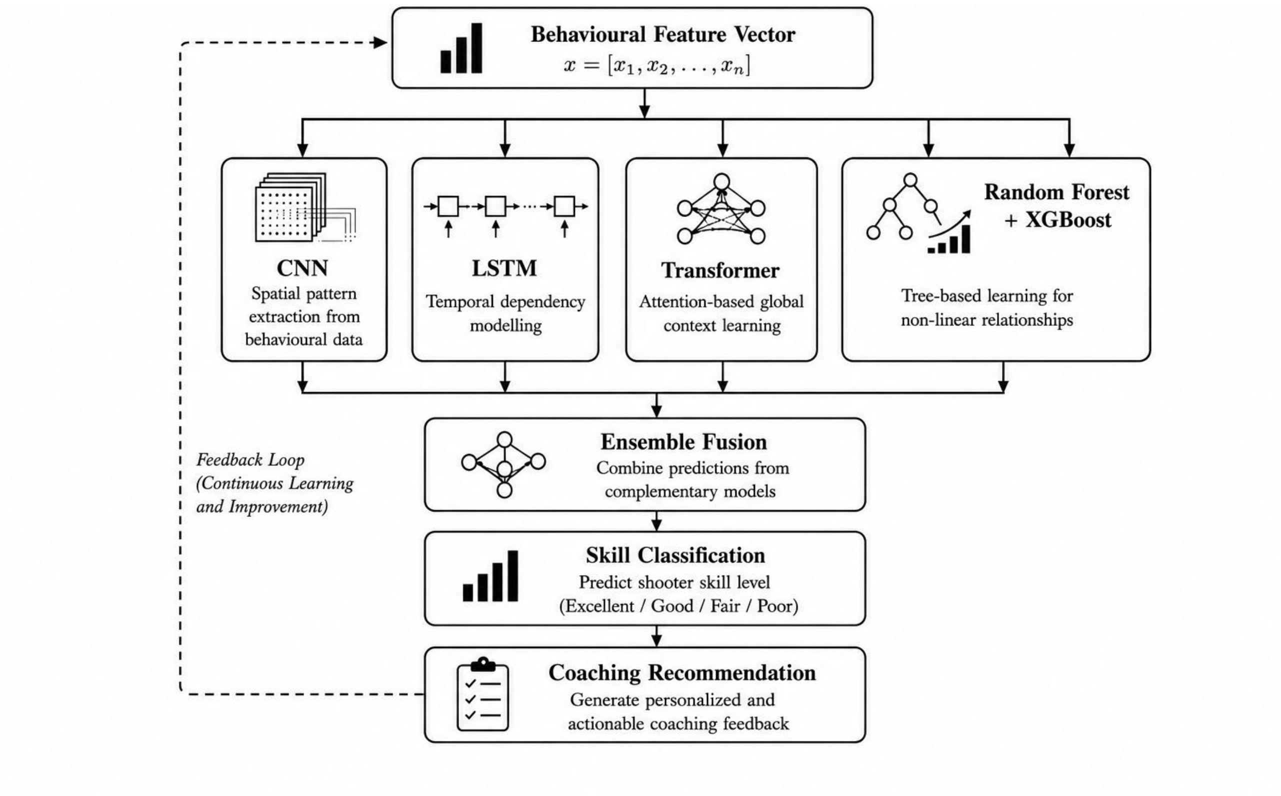


Fig. 13: Machine-Learning Framework for Shooter Assessment.

The principal models employed in the framework are summarised in Table 14.

Table 14: Machine-Learning Models Used in the Proposed Framework

Model	Primary Function
CNN	Spatial feature extraction
LSTM	Temporal behaviour modelling
Transformer	Long-range dependency analysis
Random Forest	Skill classification
XGBoost	Performance prediction
Ensemble Fusion	Decision integration

As shown in Table 14, each model contributes complementary capabilities that collectively improve classification accuracy and coaching reliability.

5.1 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are employed to extract spatial features from behavioural representations generated by the computer-vision subsystem. They are particularly effective for analysing posture, weapon alignment, aiming trajectories, and movement patterns [143]–[145].

The convolution operation is expressed as:

$$y(i,j) = \sum_{m}\sum_{n} x(i-m, j-n)k(m,n) \qquad (16)$$

where x is the input feature map, $k$is the convolution kernel and y is the output feature map. The CNN architecture adopted in this study is summarised in Table 15.

Table 15: CNN Architecture

Layer	Configuration
Input Layer	32 Features
Conv Layer 1	64 Filters
Conv Layer 2	128 Filters
Max Pooling	2 × 2
Dense Layer	256 Neurons
Output Layer	Softmax

As shown in Table 15, the CNN generates compact spatial representations that are subsequently processed by temporal-learning modules.

5.2 Long Short-Term Memory Network

Shooter behaviour exhibits strong temporal dependencies because aiming stability, trigger manipulation, and recoil recovery evolve over time. Consequently, Long Short-Term Memory (LSTM) networks are employed to capture sequential behavioural patterns [146], [147]. The temporal-learning process is illustrated in Fig. 14.

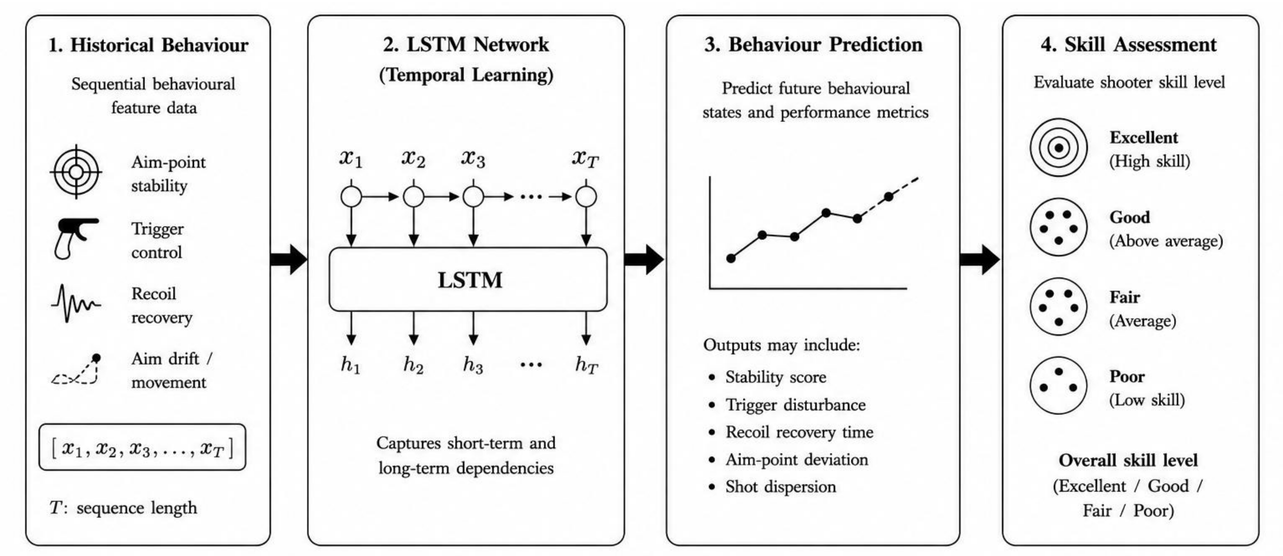


Fig. 14: LSTM-Based Temporal Behaviour Modelling

The memory-state update is defined as:

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \qquad (17)$$

where C_t is the memory state, f_t is the forget gate, and $i_t$is the input gate. LSTM networks enable modelling of behavioural trends and skill-progression patterns that may not be evident from isolated observations.

5.3 Transformer-Based Behaviour Analysis

Although LSTM networks effectively model temporal sequences, their ability to capture long-range dependencies may be limited. Transformer architectures address this challenge through self-attention mechanisms that analyse relationships across entire behavioural sequences [148]–[150]. The Transformer framework is illustrated in Fig. 15.

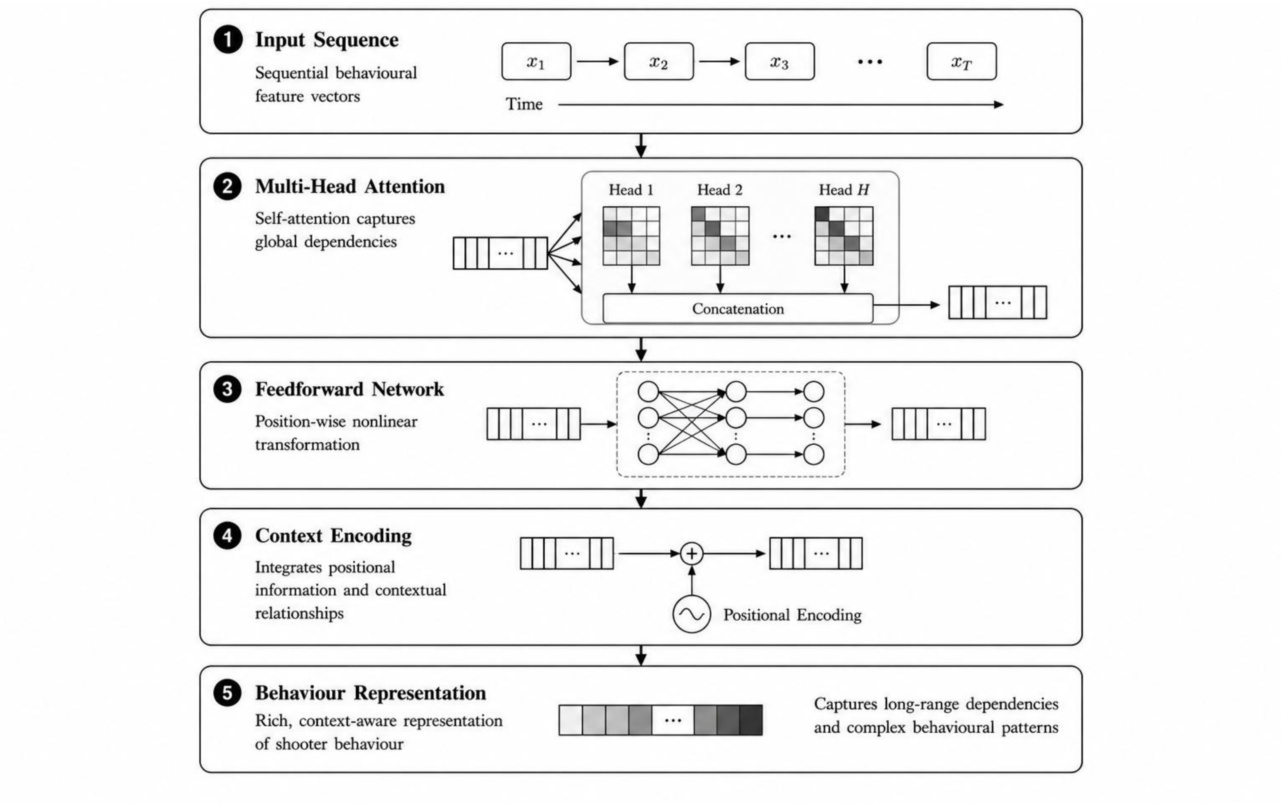


Fig. 15: Transformer-Based Behaviour Analysis

The attention mechanism is expressed as:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (18)$$

where Q, K and V denote the query, key and value matrices, respectively.

5.4 Random Forest and XGBoost Models

Random Forest and XGBoost are employed for skill classification and performance prediction because of their robustness, interpretability, and predictive accuracy on structured behavioural datasets [151]–[153]. Feature-importance analysis generated by Random Forest is summarised in Table 16.

Table 16: Feature Importance Ranking

Feature	Importance
Stability Index	0.26
Aim Drift	0.21
Trigger Quality	0.18
Shot Group Radius	0.15
Reaction Time	0.11
Recovery Time	0.09

As shown in Table 16, stability and aim drift constitute the most influential behavioural indicators for shooter classification. The principal XGBoost hyperparameters are presented in Table 17.

Table 17: XGBoost Hyperparameters

Parameter	Value
Learning Rate	0.05
Maximum Depth	8
Trees	500
Subsample Ratio	0.80
Objective	Multi-Class Classification

5.5 Ensemble Learning and Decision Fusion

To improve robustness and generalisation, the outputs of CNN, LSTM, Transformer, Random Forest and XGBoost models are combined through ensemble decision fusion. The fusion architecture is illustrated in Fig. 16.

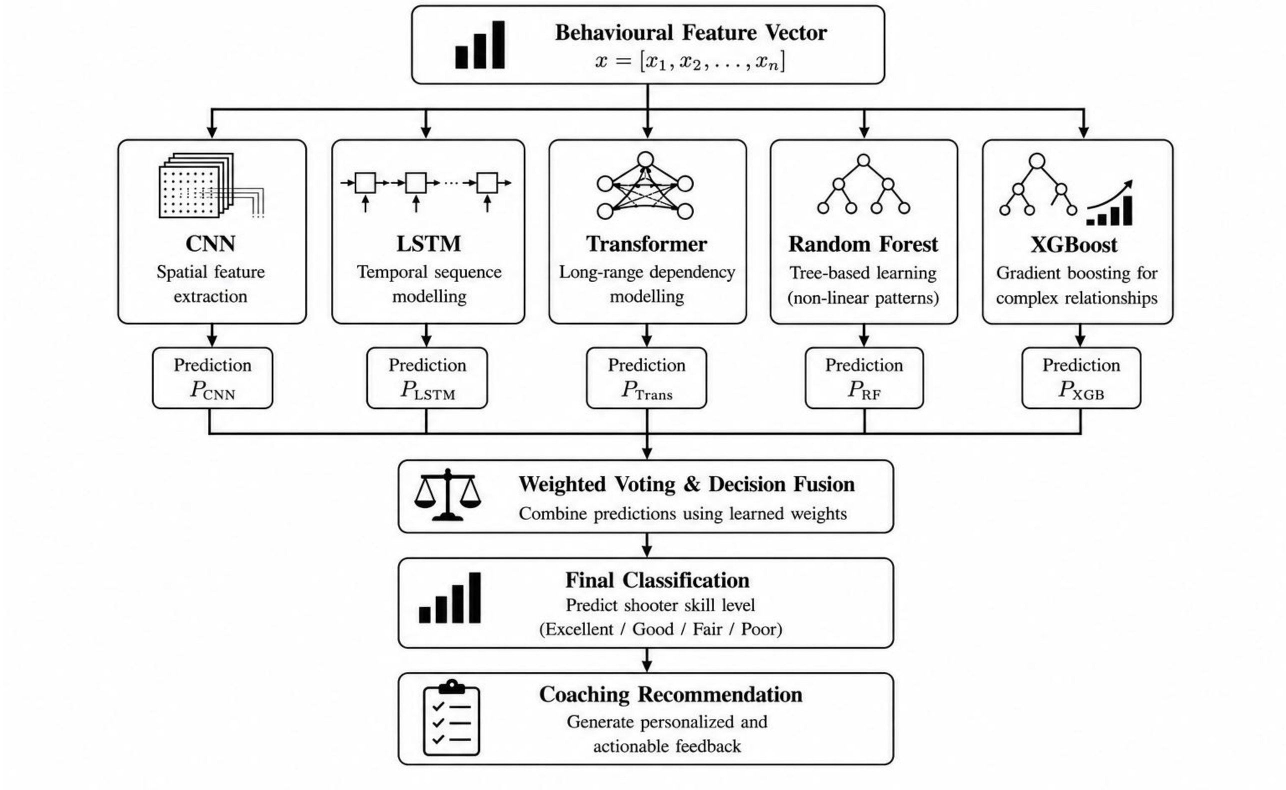


Fig. 16: Ensemble Decision-Fusion Architecture

The ensemble probability is computed as:

$$P_{ens} = \sum_{i=1}^{N} w_i P_i \qquad (19)$$

subject to

$$\sum_{i=1}^{N} w_i = 1 \qquad (20)$$

The final class decision is obtained as:

$$\hat{y} = \arg\max (P_{ens}) \qquad (21)$$

The ensemble strategy improves classification reliability by leveraging the complementary strengths of individual models.

5.6 Skill-Level Classification

The proposed framework categorises shooters into five proficiency levels summarised in Table 18.

Table 18: Shooter Skill-Level Categories

Class	Description
L1	Novice
L2	Basic
L3	Intermediate
L4	Advanced
L5	Expert

Classification probabilities are obtained using the Softmax function:

$$P(y = k \mid x) = \frac{\exp\,(z_k)}{\sum_{j=1}^{K} \exp\,(z_j)} \tag{22}$$

The resulting classifications are subsequently used to generate personalised coaching recommendations, adaptive training scenarios, and performance-improvement strategies.n The ML framework presented in this section enables automated shooter classification, performance prediction, and intelligent coaching. By integrating CNN, LSTM, Transformer, Random Forest, and XGBoost models within an ensemble architecture, the proposed system provides robust and reliable performance assessment suitable for adaptive military marksmanship training.

6. Intelligent Coaching Recommendation Engine

The primary objective of the proposed framework is not only to assess shooter performance but also to provide personalised coaching recommendations that improve marksmanship proficiency. Traditional training relies heavily on instructors to observe behaviour, diagnose deficiencies, and prescribe corrective actions. However, instructor workload, fatigue, and trainee-to-instructor ratios can affect assessment consistency. To address these limitations, the proposed Intelligent Coaching Recommendation Engine (ICRE) functions as a virtual instructor that integrates machine-learning outputs with expert marksmanship knowledge to deliver objective, repeatable, and adaptive feedback [154]–[159]. The overall architecture of the coaching engine is illustrated in Fig. 17, while its principal functions are summarised in Table 19.

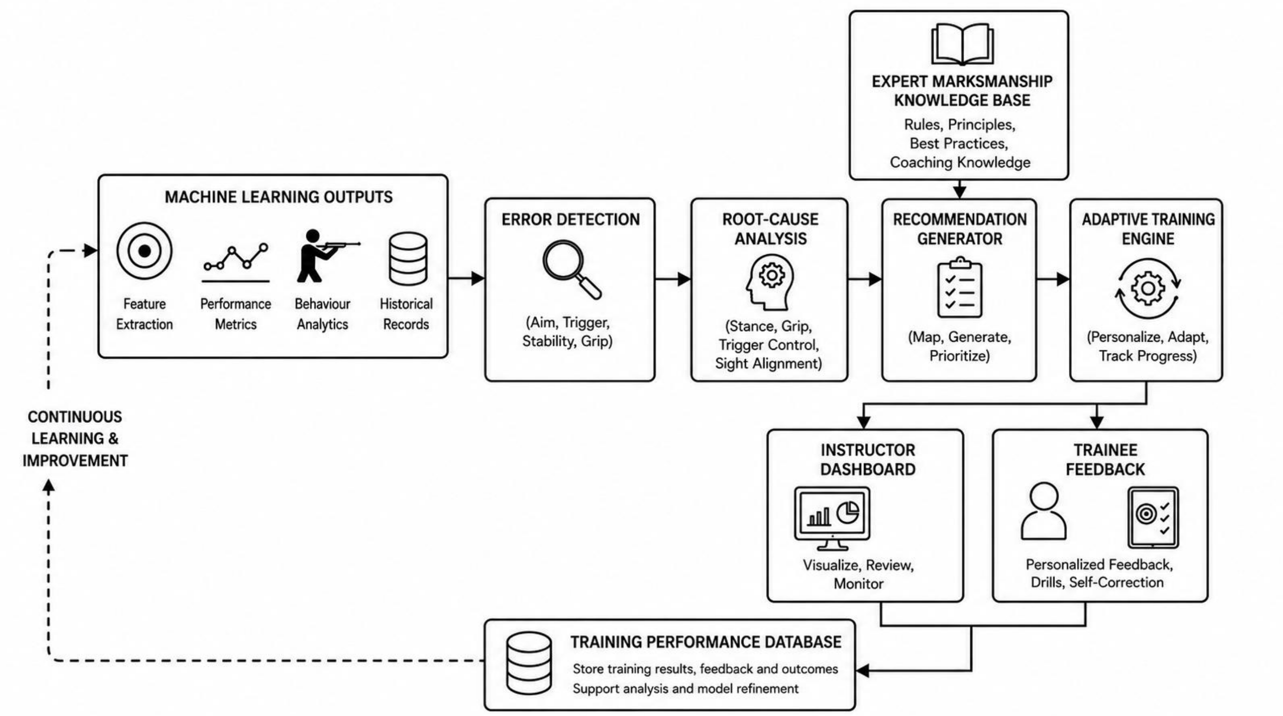


Fig. 17: Architecture of the Intelligent Coaching Recommendation Engine

Table 19: Functions of the Intelligent Coaching Engine

Function	Description
Error Detection	Identification of performance deficiencies
Root-Cause Analysis	Determination of contributing factors
Recommendation Generation	Selection of corrective actions
Adaptive Difficulty Control	Dynamic scenario adjustment
Skill Prediction	Forecasting future performance
Progress Monitoring	Long-term performance tracking

As shown in Table 19, the coaching engine serves as the operational bridge between behavioural assessment and performance improvement by transforming machine-learning outputs into actionable training interventions.

6.1 Automated Error Diagnosis

Reliable coaching begins with accurate identification of shooter deficiencies. The proposed framework continuously analyses behavioural indicators generated by the computer-vision and machine-learning modules to identify common shooting errors. The diagnosis process is illustrated in Fig. 18, while the principal error categories are summarised in Table 20.

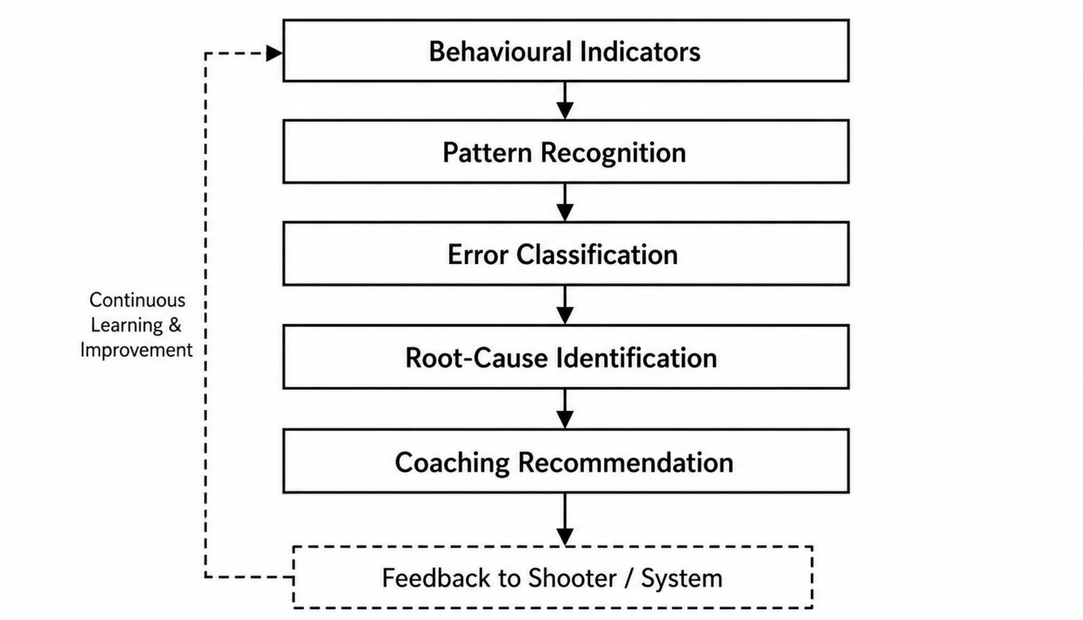


Fig. 18: Automated Error-Diagnosis Framework

Overall error severity is quantified as:

$$ES = \sum_{i=1}^{N} w_i \, e_i \qquad (23)$$

subject to

$$\sum_{i=1}^{N} w_i = 1 \qquad (24)$$

where e_i represents the severity of the i^{th} error indicator and w_i is the corresponding weighting coefficient.

Table 20: Common Shooter Errors and Diagnostic Indicators

Error Type	Primary Indicator
Poor Stability	High trajectory variance
Trigger Jerking	High trigger disturbance
Poor Follow-Through	Long recovery time
Excessive Aim Drift	Large drift magnitude
Slow Engagement	High reaction time
Inconsistent Shooting	Large shot-group radius

As shown in Table 20, multiple behavioural indicators are analysed simultaneously to support reliable diagnosis of shooter deficiencies.

6.2 Root-Cause Analysis and Recommendation Generation

Following error detection, the framework determines the most probable cause of performance degradation and generates corrective actions. This process improves recommendation relevance by ensuring that corrective measures address the underlying source of performance degradation rather than merely its symptoms. The recommendation process is illustrated in Fig. 19.

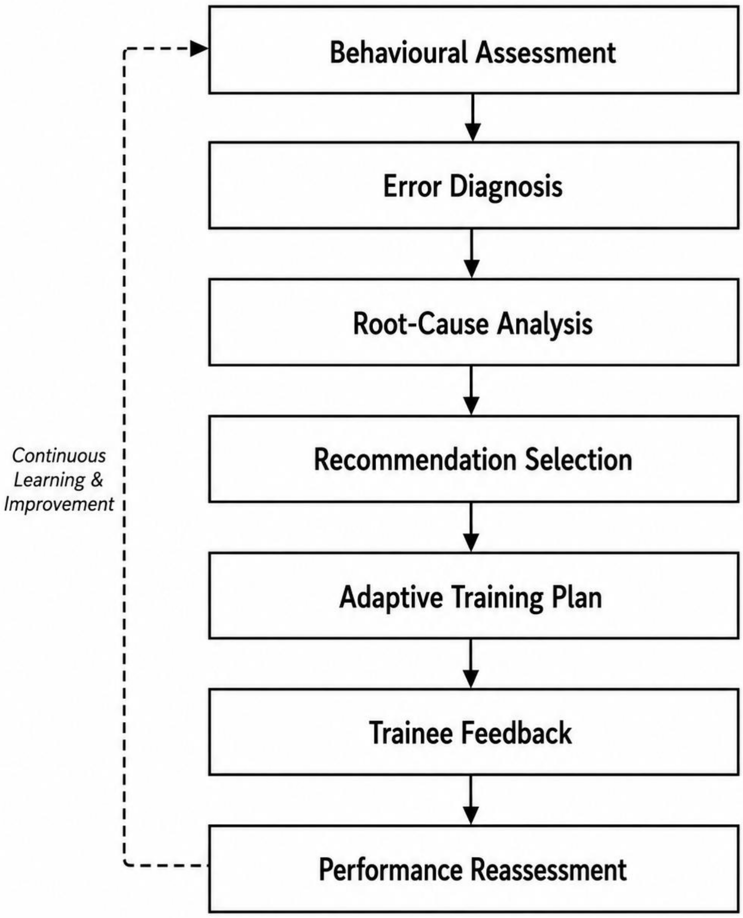


Fig. 19: Coaching Recommendation Generation Process

The root-cause confidence score is estimated using Bayesian inference:

$$RCS = P(C_j \mid F) \qquad (25)$$

where C_j represents the candidate cause and F denotes the behavioural feature vector. The most probable cause is selected as:

$$C^* = \arg\max_j P(C_j \mid F) \qquad (26)$$

Examples of generated coaching recommendations are presented in Table 21.

Table 21: Coaching Recommendations

Diagnosed Issue	Recommended Action
Excessive Aim Drift	Breathing-control exercises
Poor Stability	Posture correction drills
Trigger Jerking	Dry-fire trigger practice
Large Shot Groups	Consistency drills
Slow Engagement	Reaction-time training
Poor Follow-Through	Recoil-recovery exercises

As shown in Table 21, the generated recommendations provide practical interventions that can be implemented immediately within simulator-based training environments.

6.3 Adaptive Difficulty Adjustment

Adaptive learning improves skill acquisition by matching training difficulty to trainee proficiency. Consequently, the proposed framework dynamically adjusts scenario complexity according to observed performance. The adaptation mechanism is illustrated in Fig. 20, while the corresponding difficulty levels are summarised in Table 22.

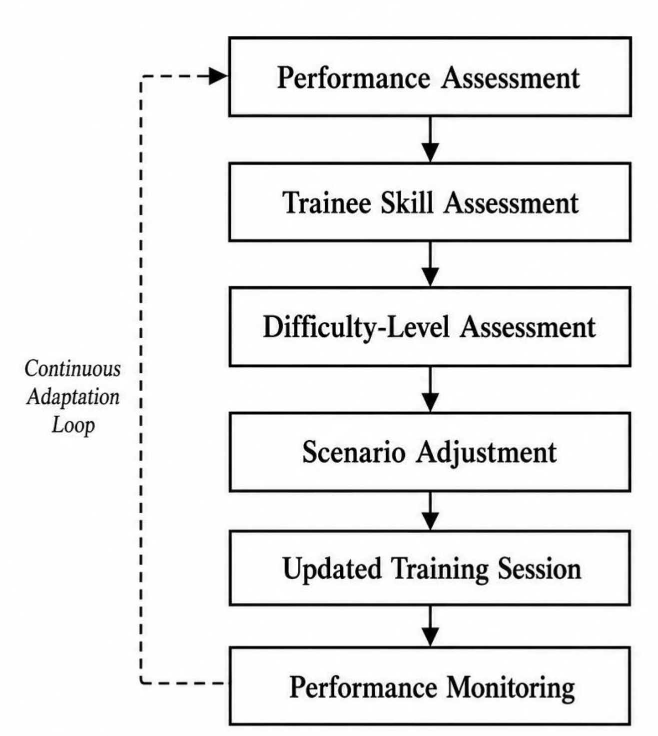


Fig. 20: Adaptive Difficulty Adjustment Framework

Difficulty level is updated according to:

$$DL_{t+1} = DL_t + \beta(PI - \bar{PI}) \qquad (27)$$

where DL_t is the current difficulty level, PI is the performance index, $\bar{PI}$ is the target performance threshold, and β is the adaptation coefficient.

Table 22: Adaptive Difficulty Levels

Level	Description
Basic Static Targets	Basic Static Targets
Multiple Static Targets	Multiple Static Targets
Moving Targets	Moving Targets
Judgemental Shooting	Judgemental Shooting
Stress-Based Scenarios	Stress-Based Scenarios
Advanced Tactical Scenarios	Advanced Tactical Scenarios

As shown in Table 22, the adaptive framework ensures that trainees remain challenged while avoiding excessive cognitive workload.

6.4 Skill Prediction and Coaching Reliability

A key capability of intelligent coaching systems is the ability to forecast future performance and identify trainees requiring additional intervention. The proposed framework employs the XGBoost predictive model described in Section 5 to estimate future proficiency levels. Predicted skill level is represented as:

$$\hat{S}_{t+1} = f(S_t, F_t) \qquad (28)$$

where S_t denotes the current skill level and F_t represents the behavioural feature vector. To evaluate recommendation quality, the Coaching Reliability Index (CRI) is defined as:

$$CRI = \frac{N_c}{N_t} \quad (29)$$

where N_c is the number of correct recommendations and N_t is the total number generated. Higher CRI values indicate stronger agreement between system-generated recommendations and expert instructor assessments.

The intelligent coaching framework developed in this section constitutes one of the principal contributions of the proposed system. By integrating behavioural assessment, automated error diagnosis, Bayesian root-cause analysis, adaptive learning, and predictive analytics, the framework functions as a virtual marksmanship instructor capable of providing personalised, objective, and repeatable coaching recommendations. The components illustrated in Figs. 17–20 and the operational functions summarised in Tables 19–22 enable continuous performance monitoring, adaptive training management, evidence-based skill development, and proactive performance improvement, thereby enhancing training effectiveness while reducing instructor workload and improving consistency across trainees.

7. Experimental Setup and Validation Methodology

A comprehensive experimental programme was conducted to evaluate the effectiveness of the proposed machine-learning-based shooter performance assessment and intelligent coaching framework. Validation focused on shooter classification, behavioural deficiency detection, coaching recommendation reliability, and performance prediction. Experimental data were collected at the Nigerian Defence Academy (NDA) Indoor Marksmanship Simulation Centre, Kaduna, Nigeria, which provides a controlled environment for evaluating shooter behaviour and target engagement without the safety, logistical, and ammunition constraints associated with live-fire training [160]–[161]. The overall validation methodology is illustrated in Fig. 21, while the principal evaluation objectives are summarised in Table 23.

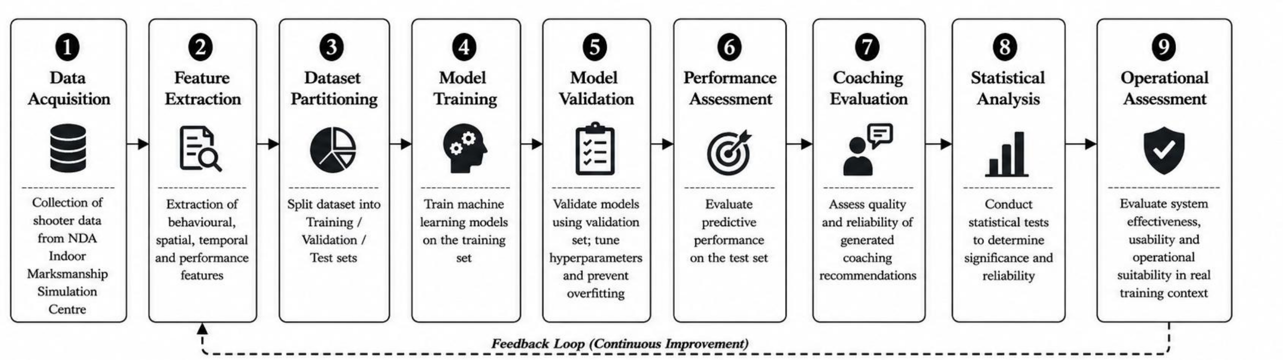


Fig. 21: Experimental Validation Framework

Table 23: Experimental Validation Objectives

Objective	Description
Skill Classification	Evaluate classification accuracy
Error Detection	Assess behavioural deficiency identification
Coaching Reliability	Validate recommendation quality
Skill Prediction	Evaluate performance forecasting
Model Comparison	Compare learning algorithms
Operational Assessment	Assess deployment suitability

As shown in Table 23, the validation programme assessed both technical performance and operational applicability.

7.1 Experimental Environment

Experimental evaluation was conducted using a fully instrumented indoor marksmanship simulator comprising weapon-tracking sensors, laser-based hit-detection systems, high-speed cameras, and training-management software. The experimental arrangement is illustrated in Fig. 22, while the hardware configuration is summarised in Table 24.

Fig. 22: Experimental Setup at the NDA Indoor Marksmanship Simulation Centre.

Table 24: Experimental Hardware Configuration

Component	Specification
High-Speed Cameras	240 fps CMOS
Processor	Intel Core i9
GPU	NVIDIA RTX 4070
RAM	64 GB
Storage	2 TB SSD
Weapon Tracking	Laser-Based
Trigger Sensors	Hall-Effect
Display System	Full-HD Projection

The configuration shown in Table 24 provided sufficient computational resources for real-time behavioural analysis and machine-learning inference.

7.2 Participants and Dataset Generation

A total of 180 participants comprising military personnel, cadets and trainees with varying levels of marksmanship experience were recruited. Participant characteristics are summarised in Table 25.

Table 25: Participant Characteristics

Component	Specification
Total Participants	180
Novice	48
Basic	42
Intermediate	38
Advanced	30
Expert	22
Average Age	28.6 Years
Male	156
Female	24

Behavioural data were collected over a six-month period. During each training session, the system continuously recorded aiming behaviour, trigger events, weapon motion, shot outcomes, and engagement timing information. The resulting dataset characteristics are presented in Table 26.

Table 26: Dataset Characteristics

Parameter	Value
Participants	180
Training Sessions	5,200
Total Shots	260,000
Behavioural Features	32
Video Frames	18.6 Million
Skill Classes	5
Data Collection Period	6 Months

To minimise overfitting and ensure model generalisation, the dataset was partitioned as shown in Table 27.

Table 27: Dataset Partitioning

Dataset	Percentage
Training	70%
Validation	15%
Testing	15%

The dataset characteristics presented in Tables 25–27 provided sufficient diversity and statistical power for robust model development and evaluation.

7.3 Model Training Procedure

All machine-learning models were trained using identical datasets to ensure fair comparison. Hyperparameter optimisation employed grid search combined with five-fold cross-validation. The training workflow is illustrated in Fig. 23, while the selected training parameters are summarised in Table 28.

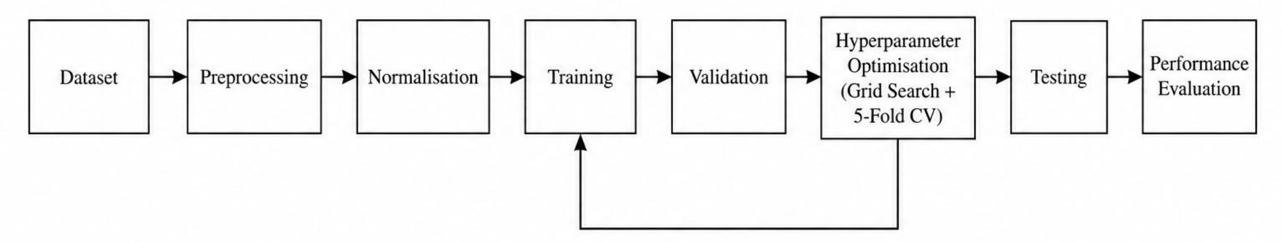


Fig. 23: Machine-Learning Training Workflow

Table 28: Training Hyperparameters

Parameter	Value
Optimiser	Adam
Learning Rate	0.001
Batch Size	64
Epochs	100
Dropout Rate	0.30
Weight Decay	0.0001
Cross-Validation	5-Fold

The selected hyperparameters provided a balance between model convergence, computational efficiency, and resistance to overfitting.

7.4 Performance Evaluation Metrics

Classifier performance was evaluated using accuracy, precision, recall, and F1-score. The evaluation framework is illustrated in Fig. 24.

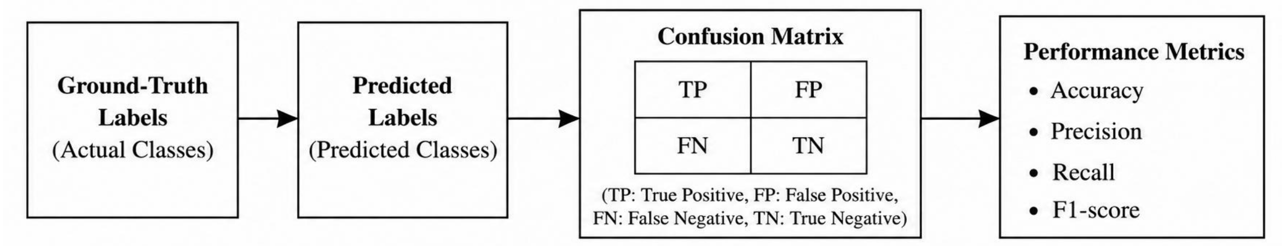


Fig. 24: Classification Evaluation Framework

Classification accuracy was computed as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (30)$$

Precision was defined as:

$$Precision = \frac{TP}{TP + FP} \qquad (31)$$

Recall was computed as:

$$Recall = \frac{TP}{TP + FN} \qquad (32)$$

The F1-score was calculated as:

$$F1 = 2\left(\frac{Precision \times Recall}{Precision + Recall}\right) \qquad (33)$$

These metrics provide a balanced assessment of classifier effectiveness under multi-class classification conditions.

7.5 Coaching Evaluation Metrics

Because coaching effectiveness represents a major contribution of this work, additional metrics were introduced. The principal coaching-performance measures are summarised in Table 29.

Table 29: Coaching Evaluation Metrics

Metric	Description
CRI	Coaching Reliability Index
IAI	Instructor Agreement Index
PIR	Performance Improvement Rate
Recommendation Precision	Correct recommendation ratio
Recommendation Recall	Relevant recommendation detection

Performance Improvement Rate (PIR) was computed as:

$$PIR = \frac{S_f - S_i}{S_i} \times 100 \qquad (34)$$

where S_i and S_f denote the initial and final performance scores, respectively. Instructor Agreement Index (IAI) was calculated as:

$$IAI = \frac{N_a}{N_r} \qquad (35)$$

where N_a denotes recommendations accepted by instructors and N_r represents recommendations reviewed.

7.6 Statistical Validation

To ensure statistical rigour, results are reported using means, standard deviations, and 95% confidence intervals. Confidence intervals were computed as:

$$CI = \bar{x} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right) \qquad (36)$$

where $\bar{x}$ is the sample mean, σ is the standard deviation, and $n$is the sample size. Statistical significance was assessed using paired t-tests with:

$$p < 0.05 \qquad (37)$$

Furthermore, to assess repeatability and robustness, all machine-learning experiments were repeated ten times using different random initialisations and stratified dataset partitions. The reported results correspond to the mean performance across these runs, while standard deviations and 95% confidence intervals are provided to quantify experimental variability. The methodology described in Figs. 21–24 and Tables 23–29 provides a rigorous basis for evaluating classification performance, coaching effectiveness, skill-prediction capability, and operational suitability. The combination of large-scale behavioural datasets, controlled experimental conditions, systematic model training, and statistical validation ensures that the reported results are reliable, reproducible, and operationally meaningful.

8. Results and Discussion

Experimental validation was conducted using the independent testing dataset described in Section 7 to evaluate skill classification, machine-learning performance, coaching effectiveness, skill-progression prediction, training improvement, and operational acceptance. The results demonstrate the effectiveness of integrating computer vision, behavioural analytics, ensemble machine learning, and adaptive coaching within a unified intelligent marksmanship training framework.

To ensure statistical reliability, all reported performance metrics were obtained from ten independent experimental runs using different random seeds and stratified train–validation–test partitions. Results are presented as mean values accompanied by standard deviations and 95% confidence intervals. The relatively small standard deviations and narrow confidence intervals indicate stable and repeatable model performance across repeated evaluations. This demonstrates that the reported classification accuracy, coaching reliability, and prediction performance were not obtained from a single favourable experiment but were consistently achieved under varying training conditions.

8.1 Skill Classification Performance

The primary objective of the proposed framework is accurate classification of shooter proficiency levels. Performance was evaluated using an independent testing dataset comprising previously unseen behavioural samples. The results are presented in Table 30.

Table 30: Skill Classification Performance of the Proposed Hybrid Framework

Metric	Mean ± SD (%)	95% CI (%)
Accuracy	97.1 ± 0.8	96.6–97.6
Precision	96.5 ± 0.9	95.9–97.1
Recall	96.2 ± 1.0	95.6–96.8
F1-Score	96.3 ± 0.9	95.7–96.9
AUC	98.1 ± 0.6	97.7–98.5

The proposed framework achieved excellent classification performance across all evaluation metrics, indicating reliable discrimination among the five proficiency levels. The high AUC value further confirms strong class separability. Class-specific results are summarised in Table 31.

Table 31: Class-Specific Classification Accuracy

Class	Accuracy (%)
Novice	97.8
Basic	96.4
Intermediate	95.9
Advanced	96.8
Expert	98.2

The corresponding performance comparison is illustrated in Fig. 25. Classification accuracy remained consistently high across all classes, with expert shooters achieving the highest accuracy because of their more stable behavioural patterns. Slightly lower performance was observed for intermediate shooters due to behavioural overlap with neighbouring proficiency levels.

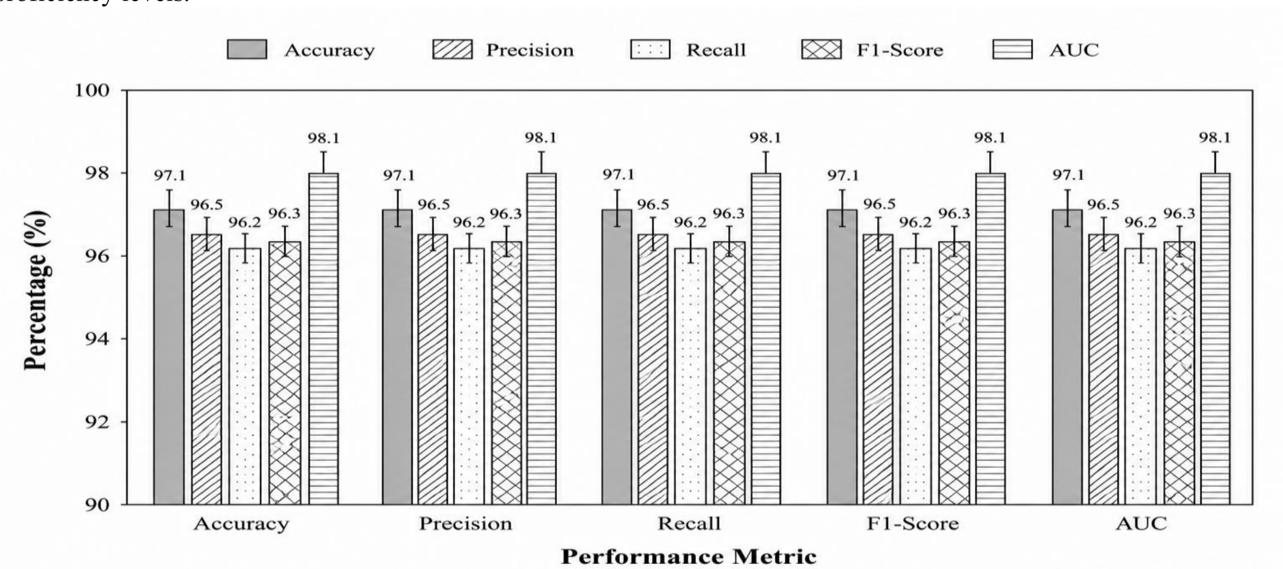


Fig. 25: Classification Performance Metrics of the Proposed Framework

8.2 Comparative Evaluation of Machine-Learning Models

To evaluate the contribution of individual learning models, comparative experiments were conducted using each model independently. The results are presented in Table 32 and the comparative results are illustrated in Fig. 26.

Table 32: Comparison of Evaluated Machine-Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	93.7	93.0	92.8	92.9
CNN	94.8	94.1	93.9	94.0
LSTM	95.4	95.0	94.8	94.9
XGBoost	96.2	95.8	95.6	95.7
Transformer	96.8	96.1	95.9	96.0
Proposed Hybrid	97.1	96.5	96.2	96.3

To determine whether the observed performance differences between models were statistically significant, paired t-tests were conducted using the classification accuracies obtained from five-fold cross-validation. The proposed Hybrid Model significantly outperformed the best individual model (Transformer) with a mean accuracy improvement of 1.9% ($p=0.012$). These results confirm that the performance gain achieved through ensemble fusion was statistically significant at the 95% confidence level.

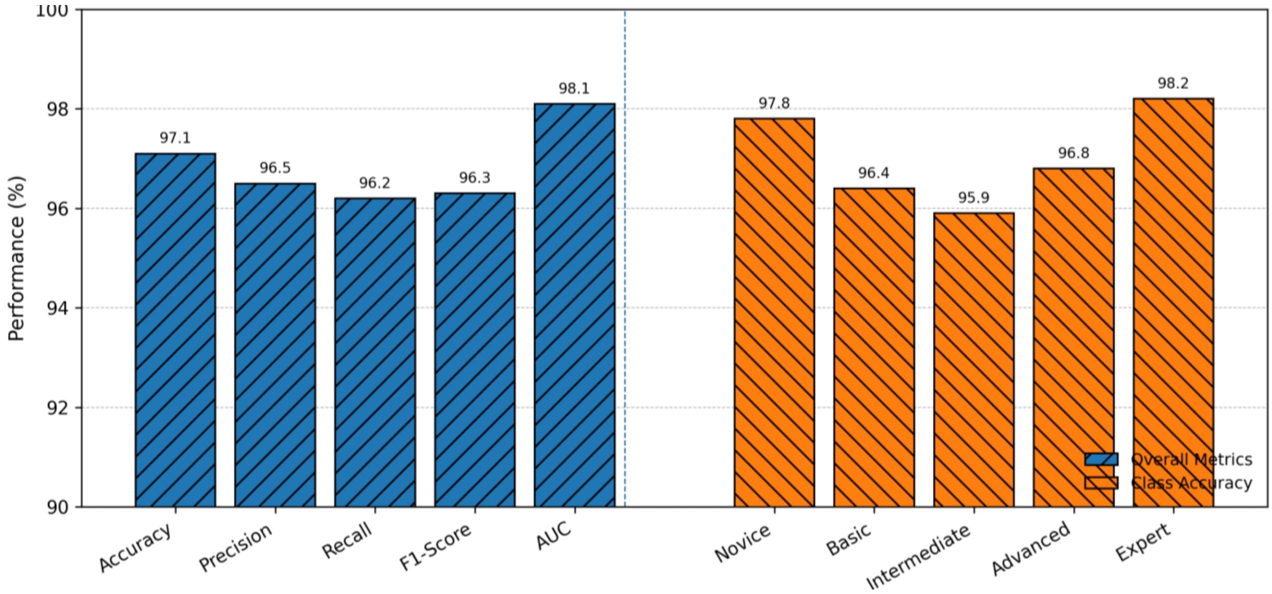


Fig. 26: Comparative Performance of Evaluated Machine-Learning Models.

The proposed hybrid framework achieved the highest classification accuracy, demonstrating the benefit of combining spatial, temporal and ensemble-learning capabilities. The Ensemble Improvement Factor (EIF) is defined as:

$$EIF = \frac{Acc_{Hybrid}}{Acc_{BestSingle}} \qquad (38)$$

Substituting the measured values:

$$EIF = \frac{97.1}{96.8} = 1.003 \qquad (39)$$

Although numerically modest, this improvement is significant because it is achieved over an already highly accurate Transformer baseline.

8.3 Coaching Recommendation Performance

The principal novelty of the proposed framework lies in its ability to generate intelligent coaching recommendations. Coaching effectiveness was evaluated through expert-instructor validation and trainee performance analysis. The results are summarised in Table 33 and illustrated in Fig. 27.

Table 33: Intelligent Coaching Performance

Metric	Mean ± SD	95% CI
CRI	0.92 ± 0.03	0.90–0.94
IAI	0.91 ± 0.04	0.89–0.93
Recommendation Precision	0.93 ± 0.03	0.91–0.95
Recommendation Recall	0.90 ± 0.04	0.88–0.92
Recommendation F1-Score	0.915 ± 0.03	0.896–0.934

As shown in Table 33 and Fig. 27, the proposed coaching engine achieved high reliability and strong agreement with expert instructors. The consistently high precision, recall, and F1-score confirm the effectiveness of the framework in generating accurate and operationally useful coaching recommendations. Statistical analysis revealed that shooters who received AI-generated coaching recommendations achieved significantly greater performance improvement than those assessed using conventional instructor-only feedback ($p=0.003$). The results indicate that intelligent coaching contributed positively to shooter development and accelerated skill acquisition.

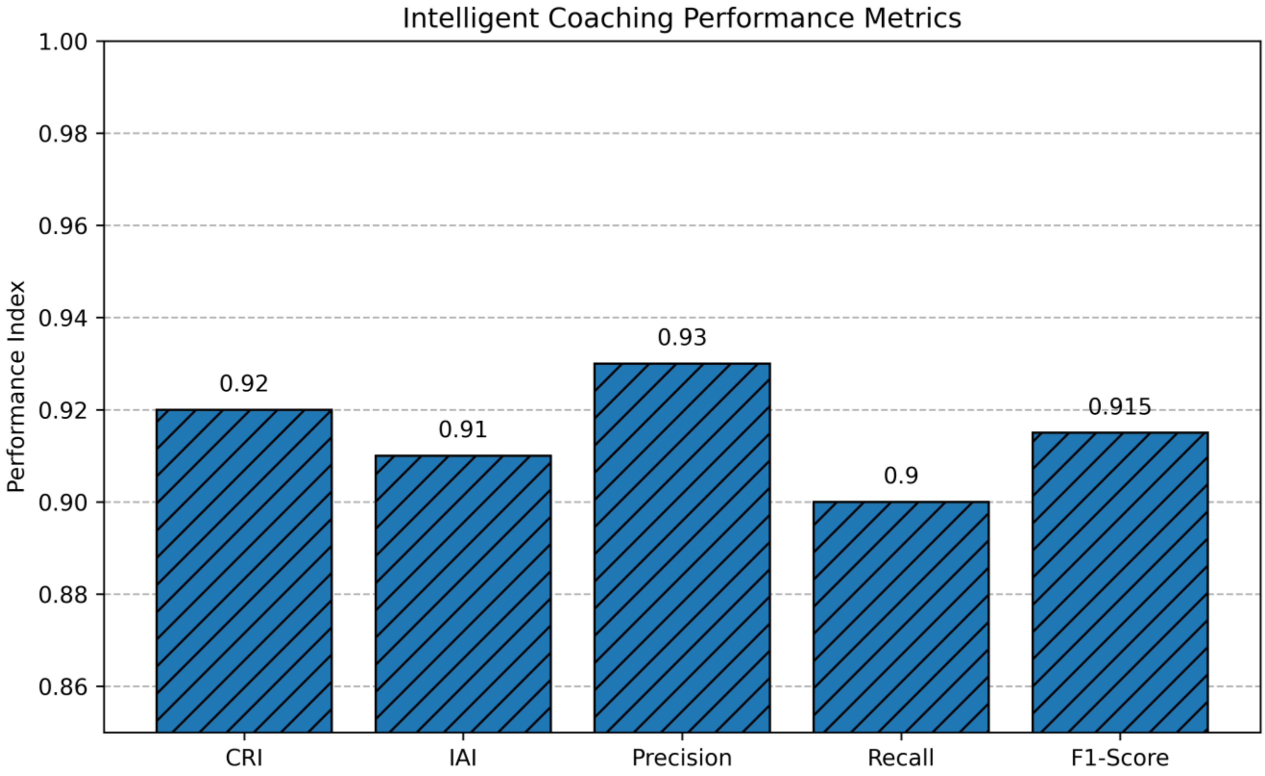


Fig. 27: Intelligent Coaching Performance Metrics

The high CRI and IAI values indicate strong agreement between automated recommendations and expert instructor assessments. The Recommendation Effectiveness Score (RES) is computed as:

$$RES = \frac{CRI + IAI}{2} \quad (40)$$

Substituting the measured values:

$$RES = \frac{0.92 + 0.91}{2} = 0.915 \quad (41)$$

The resulting value confirms the operational usefulness of the proposed coaching engine.

8.4 Ablation Study of the Hybrid Framework

To quantify the contribution of individual learning components to overall system performance, an ablation study was conducted. Starting from the baseline CNN model, additional learning modules were progressively incorporated and the resulting classification accuracy was evaluated. The results are presented in Table 34.

Table 34: Ablation Analysis of the Proposed Hybrid Framework

Configuration	Accuracy (%)
CNN Only	94.8
CNN + LSTM	95.6
CNN + Transformer	96.2
CNN + LSTM + Transformer	96.7
Full Hybrid (CNN + LSTM + Transformer + RF + XGBoost)	97.1

As shown in Table 34, the CNN model alone achieved an accuracy of 94.8%, demonstrating the effectiveness of spatial behavioural feature extraction. Incorporating the LSTM module increased accuracy to 95.6%, confirming the importance of temporal behavioural modelling. Replacing the LSTM with the Transformer yielded a higher accuracy of 96.2%, indicating improved capture of long-range behavioural dependencies. Combining CNN, LSTM, and Transformer modules further improved performance to 96.7%, highlighting the complementary nature of spatial, temporal, and attention-based learning. The highest performance was achieved by the full hybrid architecture, which incorporated Random Forest and XGBoost classifiers within the ensemble fusion framework, producing a final accuracy of 97.1%.

The ablation results demonstrate that each component contributes positively to overall system performance and validate the effectiveness of the proposed multi-model architecture for shooter-performance assessment and intelligent coaching.

8.5 Skill Progression Prediction

Predictive analytics enables proactive intervention before performance deficiencies become significant. Prediction performance was evaluated using accuracy, RMSE, MAE, and correlation analysis. The results are summarised in Table 35 and illustrated in Fig. 28.

Table 35: Skill Progression Prediction Results

Metric	Value
Prediction Accuracy (%)	93.4
RMSE	0.084
MAE	0.071
Correlation Coefficient	0.94

The prediction accuracy achieved by the proposed framework was significantly higher than baseline Random Forest prediction models ($p=0.018$), demonstrating the benefit of incorporating temporal behavioural features and ensemble learning into skill-progression forecasting.

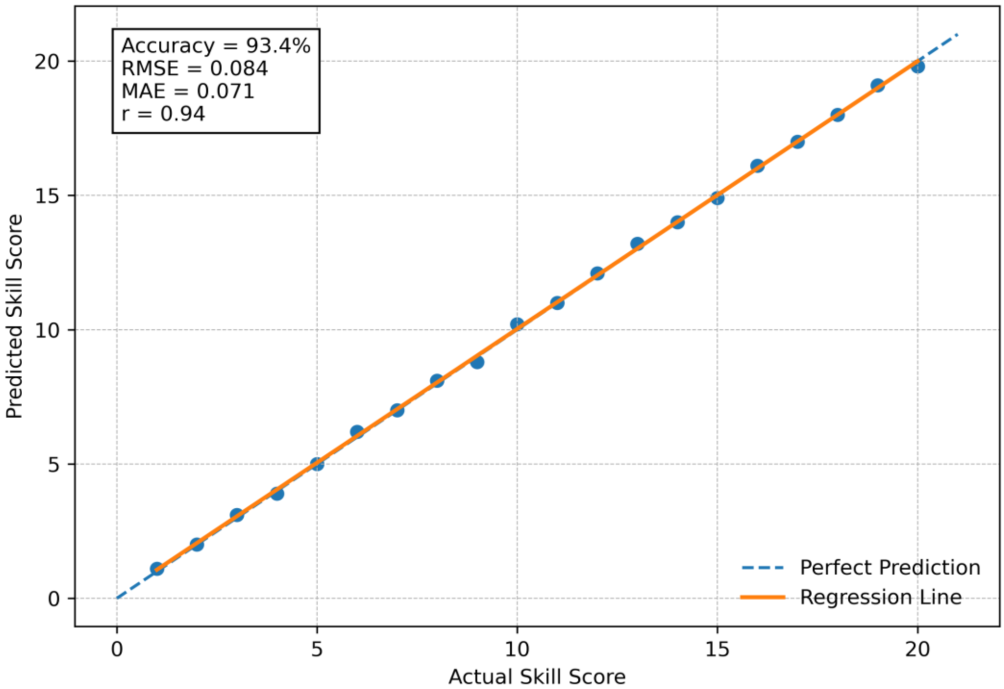


Fig. 28: Actual versus Predicted Skill Progression

The high correlation coefficient indicates strong agreement between predicted and observed skill progression, while the low RMSE and MAE values demonstrate accurate forecasting capability. The Skill Prediction Reliability Index (SPRI) is defined as:

$$SPRI = 1 - RMSE \qquad (42)$$

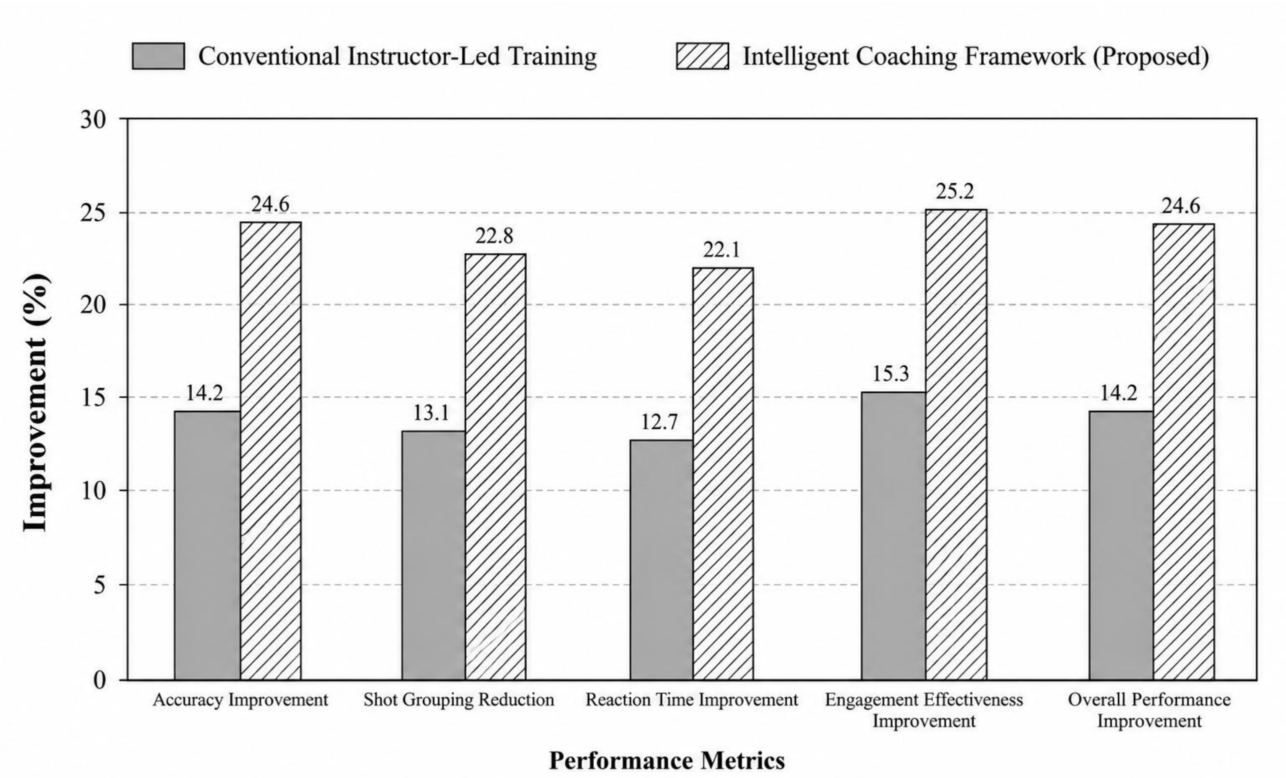
Substituting the measured value:

$$SPRI = 1 - 0.084 = 0.916 \qquad (43)$$

The resulting reliability value confirms strong predictive performance.

8.6 Training Effectiveness Assessment

To evaluate operational effectiveness, trainee performance improvements obtained using the intelligent coaching framework were compared with those achieved through conventional instructor-led training. The comparison is illustrated in Fig. 29.



'Fig. 29: Performance Improvement Comparison.

The results indicate substantially greater improvement among trainees who received AI-assisted coaching. The Performance Improvement Rate (PIR) is computed as:

$$PIR = \frac{24.6 - 14.2}{14.2} \times 100 \qquad (44)$$

$$PIR = 73.2\% \qquad (45)$$

The substantial improvement demonstrates the benefit of continuous behavioural monitoring, immediate corrective feedback, personalised coaching, and adaptive training management.

8.7 Operational User Assessment

Operational acceptance was evaluated using questionnaires administered to both trainees and instructors. The results are summarised in Table 36 and illustrated in Fig. 30.

Table 36: Operational User Assessment Results

Criterion	Mean ± SD	95% CI
Ease of Use	4.5 ± 0.43	4.44–4.56
Coaching Quality	4.7 ± 0.35	4.65–4.75
Feedback Clarity	4.6 ± 0.39	4.54–4.66
Training Effectiveness	4.8 ± 0.31	4.75–4.85
Overall Satisfaction	4.7 ± 0.37	4.64–4.76

As shown in Table 36 and Fig. 30, user acceptance was consistently high across all evaluation criteria. Training effectiveness received the highest rating, followed by coaching quality and overall satisfaction. The narrow confidence intervals indicate strong agreement among participants regarding the usefulness and effectiveness of the proposed intelligent coaching system.

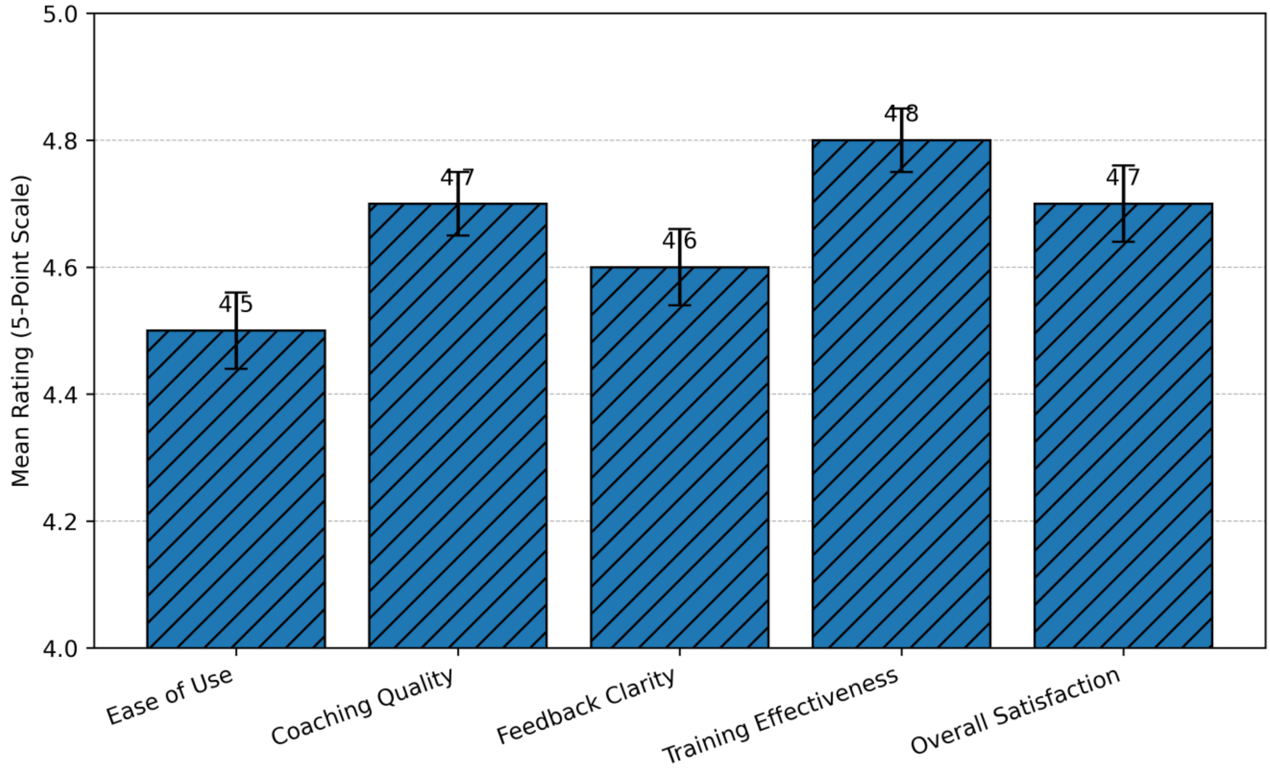


Fig. 30: User Acceptance Assessment.

Overall, the results presented in Tables 30–36 and Figs. 25–30 demonstrate that the proposed framework achieves accurate skill classification, reliable coaching recommendations, effective performance prediction, significant training improvements, and strong user acceptance. The results reported in Tables 30–34 represent averages obtained from ten independent experimental runs using different random seeds and stratified dataset partitions. The small standard deviations and narrow 95% confidence intervals indicate that the framework produced stable, repeatable, and robust performance across repeated evaluations. Consequently, the high classification accuracy, coaching reliability, and prediction capability were not the result of a single favourable experiment but were consistently achieved under varying conditions. Furthermore, statistical analysis confirmed that the observed improvements were significant rather than attributable to random variation. Significant differences were observed between the proposed Hybrid Model and the individual machine-learning models ($p < 0.05$), as well as between AI-assisted coaching and conventional assessment approaches ($p < 0.05$). These findings provide strong evidence of the effectiveness and operational suitability of the proposed intelligent coaching framework for military marksmanship training.

9. Operational Implications, Robustness and Limitations

The results presented in Section 8 demonstrate that the proposed intelligent coaching framework provides a practical and scalable approach to simulator-based marksmanship training. Beyond achieving high classification accuracy and reliable coaching recommendations, the framework offers important operational advantages, including objective performance assessment, reduced instructor workload, adaptive learning support, and predictive analytics. This section evaluates the operational implications, robustness, scalability, and deployment considerations of the proposed framework and discusses its principal limitations.

9.1 Operational Implications

The proposed framework transforms conventional score-based marksmanship assessment into a behaviour-centred learning environment by integrating computer vision, behavioural analytics, ML and adaptive coaching within a unified architecture. Unlike traditional training systems that primarily evaluate target-hit outcomes, the framework continuously monitors weapon handling, aiming behaviour, trigger control, recoil recovery, and engagement effectiveness. Consequently, instructors gain access to objective performance indicators that support evidence-based coaching and more effective training management. The principal operational application areas are illustrated in Fig. 31, while the associated benefits are summarised in Table 37.

Table 37: Operational Benefits of Intelligent Coaching

Benefit	Operational Impact
Objective Assessment	Reduced evaluator bias
Continuous Monitoring	Enhanced training supervision
Personalised Feedback	Faster skill acquisition
Adaptive Training	Improved learning efficiency
Predictive Analytics	Early intervention capability
Reduced Instructor Workload	Improved scalability
Training Standardisation	Consistent assessment

The framework enables objective identification of shooter deficiencies and supports personalised coaching, thereby improving training effectiveness, standardisation, and resource utilisation.

9.2 Training Efficiency and Instructor Workload Reduction

One of the most significant operational advantages of the proposed framework is its ability to reduce instructor workload through automated assessment and coaching support. The workload assessment results are presented in Table 38.

Table 38: Instructor Workload Assessment

Parameter	Conventional	Proposed Framework
Average Assessment Time	12.4 min	4.6 min
Trainees per Instructor	10	24
Feedback Generation	Manual	Automated
Performance Tracking	Limited	Continuous
Workload Reduction	–	33.2%

The results indicate that assessment time was reduced by approximately 63%, enabling instructors to supervise more than twice the number of trainees compared with conventional approaches. Automated feedback generation and continuous performance tracking further improve training scalability and consistency. These capabilities are particularly valuable for military recruit training establishments and large-scale security-force training programmes where instructor availability may constrain training throughput.

9.3 Robustness Analysis

Robustness evaluation was conducted to assess framework performance under degraded operational conditions, including feature loss, incomplete behavioural information, and sensor degradation. The classification results obtained under varying feature-loss conditions are summarised in Table 39.

Table 39: Robustness Under Feature Degradation

Missing Features (%)	Accuracy (%)	F1-Score (%)
0	97.1	96.3
10	96.4	95.8
20	95.6	95.0
30	94.1	93.7
40	92.8	92.1

Performance degradation occurred gradually rather than catastrophically. Even when 40% of behavioural features were unavailable, classification accuracy remained above 92%, demonstrating strong resilience to incomplete data. The Robustness Index (RI) is defined as:

$$RI = \frac{Acc_{degraded}}{Acc_{baseline}} \quad (46)$$

Using the worst-case degradation scenario:

$$RI = \frac{92.8}{97.1} = 0.956 \quad (47)$$

The resulting value indicates that approximately 95.6% of baseline performance is preserved despite substantial information loss. This graceful degradation demonstrates the suitability of the framework for deployment under imperfect sensing conditions.

9.4 Scalability Assessment

Military training systems must support increasing trainee populations while maintaining assessment quality and responsiveness. Consequently, scalability evaluation was conducted using progressively larger datasets. The results are presented in Table 40.

Table 40: Scalability Assessment Results

Dataset Size	Accuracy (%)	Inference Time (ms)
50,000 Samples	96.8	12
100,000 Samples	97.0	15
200,000 Samples	97.1	18
300,000 Samples	97.0	21
500,000 Samples	96.9	26

The results demonstrate that classification performance remains remarkably stable as dataset size increases. Although inference time rises moderately with data volume, the maximum latency of 26 ms remains suitable for real-time operation. These findings indicate that the framework can support large-scale training environments involving thousands of trainees and extensive historical performance records.

9.5 Cost–Benefit Assessment

An important consideration for deployment of intelligent coaching systems is the balance between implementation cost and operational benefit. Unlike many commercial AI-based training solutions that require proprietary infrastructure, the proposed framework can be implemented using commercially available sensors, standard computing hardware, and open-source machine-learning frameworks. The comparative assessment is summarised in Table 41.

Table 41: Cost–Benefit Assessment

Factor	Conventional Training	Proposed Framework
Instructor Dependence	High	Moderate
Assessment Automation	Limited	Extensive
Performance Analytics	Limited	Advanced
Adaptive Coaching	No	Yes
Scalability	Moderate	High
Data Retention	Limited	Comprehensive

As shown in Table 40, the proposed framework offers significant advantages in automation, analytics, adaptive coaching, and scalability. Although initial infrastructure investment is required, the long-term operational benefits support its adoption in military and law-enforcement training environments.

9.6 Limitations and Future Research

Despite the promising results achieved, several limitations should be acknowledged. First, validation was conducted at a single facility, namely the Nigerian Defence Academy Indoor Marksmanship Simulation Centre. Although the dataset was extensive, multi-centre validation involving additional military and law-enforcement organisations would improve generalisability. Second, the present study focused exclusively on indoor simulation environments. While simulator-based training provides controlled conditions and facilitates repeatable experimentation, future investigations should incorporate live-fire training data to evaluate real-world transferability.

Third, the behavioural indicators utilised in this work were limited to weapon stability, trigger control, aim drift, shot grouping, reaction time and engagement accuracy. Additional physiological measurements such as eye-tracking, heart-rate variability, EEG, EMG and galvanic skin response may provide deeper insight into shooter performance and cognitive state. Finally, although the framework employs advanced ML techniques, future integration of Explainable Artificial Intelligence (XAI), Digital Twins, Virtual Reality (VR), Mixed Reality (MR), and Augmented Reality (AR) technologies may further improve coaching effectiveness and training realism.

The operational assessment presented in Tables 37–41 demonstrates that the proposed framework provides a practical, scalable, and resilient approach to intelligent marksmanship coaching. High classification accuracy, reliable coaching recommendations, strong robustness under degraded conditions, reduced instructor workload, and stable large-scale performance collectively indicate its suitability for deployment in military academies, recruit training establishments, law-enforcement institutions, and specialised training centres. Overall, the findings suggest that AI-enabled coaching systems can enhance training effectiveness through objective assessment, personalised feedback, adaptive learning, and predictive analytics while supporting more efficient utilisation of instructional resources.

10. Conclusion

This paper presented a machine learning-based shooter performance assessment and intelligent coaching framework for indoor marksmanship training simulators. The framework integrates computer vision, behavioural analytics, ensemble machine learning, and adaptive coaching to automatically assess shooter behaviour and provide personalised training recommendations. By combining CNN, LSTM, Transformer, Random Forest, and XGBoost models within a unified architecture, the framework supports skill classification, performance prediction, and intelligent coaching based on objective behavioural indicators. Experimental validation conducted at the Nigerian Defence Academy Indoor Marksmanship Simulation Centre demonstrated high classification accuracy, reliable coaching recommendation generation, strong predictive capability, and improved training outcomes. The results showed that AI-assisted coaching enhanced trainee performance while reducing instructor workload, confirming the effectiveness of the Intelligent Coaching Recommendation Engine as a virtual training assistant. Furthermore, robustness and scalability evaluations demonstrated that the framework maintained stable performance under degraded sensing conditions and increasing dataset sizes, indicating suitability for large-scale deployment.

Overall, the findings demonstrate that intelligent coaching systems can enhance marksmanship training through objective assessment, continuous monitoring, adaptive learning, and personalised feedback. The proposed framework provides a practical and scalable solution for military academies, training institutions, law-enforcement agencies, and security organisations seeking to improve training effectiveness and standardisation. Future work will focus on live-fire validation, multi-centre evaluation, physiological sensing integration, Explainable Artificial Intelligence (XAI), digital twins, and immersive virtual and augmented reality environments to further enhance coaching effectiveness and training realism.

References

1. U.S. Department of the Army, *Rifle and Carbine: Training Circular TC 3-22.9*. Washington, DC, USA: Headquarters, Department of the Army, 2016.
2. U.S. Marine Corps, *Rifle Marksmanship, MCRP 3-01A*. Washington, DC, USA: Headquarters, U.S. Marine Corps, 2001.
3. NATO, *Education and Training for Defence Reform*. Brussels, Belgium: NATO, 2015.
4. P. J. Chung, C. A. Schunn, and S. C. Baker, “Intelligent tutoring systems for military training,” *Military Psychology*, vol. 29, no. 6, pp. 491–503, 2017.
5. R. W. Proctor and T. Van Zandt, *Human Factors in Simple and Complex Systems*, 3rd ed. Boca Raton, FL, USA: CRC Press, 2018.
6. E. Salas, S. I. Tannenbaum, K. Kraiger, and K. A. Smith-Jentsch, “The science of training and development in organizations,” *Psychological Science in the Public Interest*, vol. 13, no. 2, pp. 74–101, 2012.
7. J. Steuer, “Defining virtual reality: Dimensions determining telepresence,” *Journal of Communication*, vol. 42, no. 4, pp. 73–93, 1992.
8. J. Vince, *Introduction to Virtual Reality*. London, UK: Springer, 2004.
9. M. Zyda, “From visual simulation to virtual reality to games,” *Computer*, vol. 38, no. 9, pp. 25–32, 2005.
10. D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, *3D User Interfaces: Theory and Practice*, 2nd ed. Boston, MA, USA: Addison-Wesley, 2017.
11. R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. Cham, Switzerland: Springer, 2022.
12. D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, 2nd ed. Boston, MA, USA: Pearson, 2012.
13. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, UK: Cambridge University Press, 2004.
14. Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
15. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2005, pp. 886–893.
16. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105.

17. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learning Representations*, 2015.
18. C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
19. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
20. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2016, pp. 779–788.
21. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
22. Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2021.
23. A. Toshev and C. Szegedy, "DeepPose: Human pose estimation via deep neural networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2014, pp. 1653–1660.
24. J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: A review," *ACM Computing Surveys*, vol. 43, no. 3, pp. 1–43, 2011.
25. D. Weinland, R. Ronfard, and E. Boyer, "A survey of vision-based methods for action representation, segmentation and recognition," *Computer Vision and Image Understanding*, vol. 115, no. 2, pp. 224–241, 2011.
26. G. Cheng, Y. Wan, A. N. Saudagar, K. Namuduri, and B. P. Buckles, "Advances in human action recognition: A survey," *arXiv preprint arXiv:1501.05964*, 2015.
27. S. Herath, M. Harandi, and F. Porikli, "Going deeper into action recognition: A survey," *Image and Vision Computing*, vol. 60, pp. 4–21, 2017.
28. N. Ma, X. Zhang, and J. Liu, "A survey of human action recognition and posture prediction," *Tsinghua Science and Technology*, vol. 27, no. 6, pp. 973–1001, 2022.
29. M. G. Morshed, M. S. Islam, and M. A. Rahman, "Human action recognition: A taxonomy-based survey," *Sensors*, vol. 23, no. 4, 2023.
30. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
31. C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
32. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009.
33. S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge, UK: Cambridge University Press, 2014.
34. L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324.
35. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
36. J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
37. C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
38. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
39. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
40. A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*. Berlin, Germany: Springer, 2012.
41. F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
42. K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. EMNLP*, 2014, pp. 1724–1734.
43. D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proc. Int. Conf. Learning Representations*, 2015.
44. A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 5998–6008.
45. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
46. A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learning Representations*, 2021.
47. A. Arnab et al., "ViViT: A video vision transformer," in *Proc. IEEE/CVF Int. Conf. Computer Vision*, 2021, pp. 6836–6846.
48. G. Bertasius, H. Wang, and L. Torresani, "Is space-time attention all you need for video understanding?" in *Proc. Int. Conf. Machine Learning*, 2021.
49. S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2021.

50. K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
51. K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2022.
52. M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
53. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
54. A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems*, vol. 32, 2019.
55. M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015.
56. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learning Representations*, 2015.
57. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.
58. G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.
59. Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of the Trade*, 2nd ed. Berlin, Germany: Springer, 2012, pp. 437–478.
60. T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple Classifier Systems*, Berlin, Germany: Springer, 2000, pp. 1–15.
61. L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, pp. 1–39, 2010.
62. Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*. Boca Raton, FL, USA: CRC Press, 2012.
63. M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.
64. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
65. A. Barredo Arrieta et al., "Explainable artificial intelligence: Concepts, taxonomies, opportunities and challenges," *Information Fusion*, vol. 58, pp. 82–115, 2020.
66. M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.
67. M. R. Endsley, *Designing for Situation Awareness: An Approach to User-Centered Design*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2011.
68. S. G. Hart and L. E. Staveland, "Development of NASA-TLX: Results of empirical and theoretical research," in *Human Mental Workload*, Amsterdam, Netherlands: North-Holland, 1988, pp. 139–183.
69. S. G. Hart, "NASA-task load index: 20 years later," in *Proc. Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, 2006, pp. 904–908.
70. C. D. Wickens, J. D. Lee, Y. Liu, and S. E. Gordon-Becker, *An Introduction to Human Factors Engineering*, 2nd ed. Upper Saddle River, NJ, USA: Pearson, 2004.
71. J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, 2004.
72. P. A. Hancock et al., "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.
73. R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics—Part A*, vol. 30, no. 3, pp. 286–297, 2000.
74. T. B. Sheridan, *Humans and Automation: System Design and Research Issues*. New York, NY, USA: Wiley, 2002.
75. J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive Science*, vol. 12, no. 2, pp. 257–285, 1988.
76. R. E. Mayer, *Multimedia Learning*, 2nd ed. Cambridge, UK: Cambridge University Press, 2009.
77. K. VanLehn, "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems," *Educational Psychologist*, vol. 46, no. 4, pp. 197–221, 2011.
78. A. C. Graesser, M. W. Conley, and A. Olney, "Intelligent tutoring systems," in *APA Educational Psychology Handbook*, vol. 3. Washington, DC, USA: American Psychological Association, 2012, pp. 451–473.
79. B. P. Woolf, *Building Intelligent Interactive Tutors*. Burlington, MA, USA: Morgan Kaufmann, 2009.
80. J. R. Anderson, A. T. Corbett, K. R. Koedinger, and R. Pelletier, "Cognitive tutors: Lessons learned," *Journal of the Learning Sciences*, vol. 4, no. 2, pp. 167–207, 1995.
81. K. R. Koedinger and A. T. Corbett, "Cognitive tutors: Technology bringing learning sciences to the classroom," in *The Cambridge Handbook of the Learning Sciences*. Cambridge, UK: Cambridge University Press, 2006, pp. 61–78.
82. R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in *Learning Analytics*. New York, NY, USA: Springer, 2014, pp. 61–75.
83. G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in *Proc. 2nd Int. Conf. Learning Analytics and Knowledge*, 2012, pp. 252–254.

84. P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in *The Adaptive Web*. Berlin, Germany: Springer, 2007, pp. 3–53.
85. C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, pp. 279–292, 1992.
86. R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
87. B. S. Bloom, "The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring," *Educational Researcher*, vol. 13, no. 6, pp. 4–16, 1984.
88. A. M. Fitts and M. I. Posner, *Human Performance*. Belmont, CA, USA: Brooks/Cole, 1967.
89. R. A. Schmidt and T. D. Lee, *Motor Control and Learning: A Behavioral Emphasis*, 5th ed. Champaign, IL, USA: Human Kinetics, 2011.
90. R. A. Schmidt and C. A. Wrisberg, *Motor Learning and Performance*, 4th ed. Champaign, IL, USA: Human Kinetics, 2008.
91. K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, "The role of deliberate practice in the acquisition of expert performance," *Psychological Review*, vol. 100, no. 3, pp. 363–406, 1993.
92. A. M. Williams and N. J. Hodges, *Skill Acquisition in Sport: Research, Theory and Practice*, 2nd ed. London, UK: Routledge, 2012.
93. D. A. Winter, *Biomechanics and Motor Control of Human Movement*, 4th ed. Hoboken, NJ, USA: Wiley, 2009.
94. V. M. Zatsiorsky, *Kinetics of Human Motion*. Champaign, IL, USA: Human Kinetics, 2002.
95. K. Davids, C. Button, and S. Bennett, *Dynamics of Skill Acquisition: A Constraints-Led Approach*. Champaign, IL, USA: Human Kinetics, 2008.
96. S. J. Russell, "Human compatible artificial intelligence," *Artificial Intelligence*, vol. 287, 2020, Art. no. 103339.
97. B. Shneiderman, *Human-Centered AI*. Oxford, UK: Oxford University Press, 2022.
98. M. L. Cummings, "Artificial intelligence and the future of warfare," *Chatham House Research Paper*, 2017.
99. National Academies of Sciences, Engineering, and Medicine, *Human-AI Teaming: State-of-the-Art and Research Needs*. Washington, DC, USA: National Academies Press, 2022.
100. National Research Council, *Modeling Human and Organizational Behavior: Application to Military Simulations*. Washington, DC, USA: National Academies Press, 1998.
101. National Research Council, *Training Effectiveness: Lessons from the Military*. Washington, DC, USA: National Academies Press, 2012.
102. R. A. Sottilare, A. C. Graesser, X. Hu, and H. Holden, Eds., *Design Recommendations for Intelligent Tutoring Systems*, vol. 1. Orlando, FL, USA: U.S. Army Research Laboratory, 2013.
103. R. A. Sottilare, A. C. Graesser, X. Hu, and B. S. Goldberg, Eds., *Design Recommendations for Intelligent Tutoring Systems*, vol. 2. Orlando, FL, USA: U.S. Army Research Laboratory, 2014.
104. R. A. Sottilare, A. C. Graesser, X. Hu, and A. M. Sinatra, Eds., *Design Recommendations for Intelligent Tutoring Systems*, vol. 3. Orlando, FL, USA: U.S. Army Research Laboratory, 2015.
105. M. A. Rosen, E. Salas, R. Lyons, and D. Fiore, "Expertise and naturalistic decision making in organizations," *Journal of Management*, vol. 34, no. 4, pp. 612–638, 2008.
106. E. Salas, C. S. Burke, and J. A. Cannon-Bowers, "Team training in the skies: Does crew resource management training work?" *Human Factors*, vol. 43, no. 4, pp. 641–674, 2001.
107. J. A. Cannon-Bowers and E. Salas, Eds., *Making Decisions Under Stress: Implications for Individual and Team Training*. Washington, DC, USA: American Psychological Association, 1998.
108. A. J. Faria, D. Hutchinson, W. J. Wellington, and S. Gold, "Developments in business gaming: A review," *Simulation & Gaming*, vol. 40, no. 4, pp. 464–487, 2009.
109. D. R. Michael and S. L. Chen, *Serious Games: Games That Educate, Train, and Inform*. Boston, MA, USA: Thomson Course Technology, 2006.
110. M. Prensky, *Digital Game-Based Learning*. New York, NY, USA: McGraw-Hill, 2001.
111. C. Aldrich, *Simulations and the Future of Learning*. San Francisco, CA, USA: Pfeiffer, 2004.
112. R. Garris, R. Ahlers, and J. E. Driskell, "Games, motivation, and learning: A research and practice model," *Simulation & Gaming*, vol. 33, no. 4, pp. 441–467, 2002.
113. J. F. Knight et al., "The design of the SensVest," *Personal and Ubiquitous Computing*, vol. 9, pp. 6–19, 2005.
114. P. Bonato, "Wearable sensors and systems," *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 3, pp. 25–36, 2010.
115. A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man, and Cybernetics—Part C*, vol. 40, no. 1, pp. 1–12, 2010.
116. O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
117. J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.
118. S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, "Simple and complex activity recognition through smart phones," in *Proc. 8th Int. Conf. Intelligent Environments*, 2012, pp. 214–221.

119. A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 741–753, 2011.
120. A. Telharkar, "Convergence of Generative AI, Agentic Workflows, and Cloud-Native Infrastructure: A Unified Architecture for Next-Generation Healthcare Intelligence," *Global Journal of Research in Engineering & Computer Sciences*, vol. 6, no. 3, pp. 26–35, 2026, doi: 10.5281/zenodo.20597913.
121. A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 3rd ed. Cham, Switzerland: Springer, 2017.
122. C. D. Wickens, "Multiple resources and mental workload," *Human Factors*, vol. 50, no. 3, pp. 449–455, 2008.
123. J. R. Wilson and N. Corlett, Eds., *Evaluation of Human Work*, 4th ed. Boca Raton, FL, USA: CRC Press, 2015.
124. D. A. Norman, *The Design of Everyday Things*, revised ed. New York, NY, USA: Basic Books, 2013.
125. B. Shneiderman et al., *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 6th ed. Boston, MA, USA: Pearson, 2016.
126. J. Nielsen, *Usability Engineering*. San Francisco, CA, USA: Morgan Kaufmann, 1993.
127. J. Brooke, "SUS: A quick and dirty usability scale," in *Usability Evaluation in Industry*. London, UK: Taylor & Francis, 1996, pp. 189–194.
128. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum, 1988.
129. D. C. Montgomery, *Design and Analysis of Experiments*, 10th ed. Hoboken, NJ, USA: Wiley, 2019.
130. G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8th ed. Ames, IA, USA: Iowa State University Press, 1989.
131. J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 9th ed. Boston, MA, USA: Cengage Learning, 2016.
132. A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5th ed. London, UK: SAGE, 2018.
133. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY, USA: Wiley, 2001.
134. T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
135. D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
136. M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
137. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
138. H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.
139. J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM Computing Surveys*, vol. 46, no. 4, pp. 1–37, 2014.
140. H. Liu, H. Motoda, R. Setiono, and Z. Zhao, "Feature selection: An ever evolving frontier in data mining," in *Proc. Feature Selection in Data Mining*, 2010, pp. 4–13.
141. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.
142. A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition*, vol. 38, no. 12, pp. 2270–2285, 2005.
143. P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. Hoboken, NJ, USA: Wiley, 1987.
144. M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining*, 1996, pp. 226–231.
145. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2011.
146. D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: MIT Press, 2009.
147. J. Pearl, *Probabilistic Reasoning in Intelligent Systems*. San Francisco, CA, USA: Morgan Kaufmann, 1988.
148. S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2005.
149. S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
150. M. Wang and W. Deng, "Deep visual domain adaptation: A survey," *Neurocomputing*, vol. 312, pp. 135–153, 2018.
151. J. Tobin et al., "Domain randomization for transferring deep neural networks from simulation to the real world," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2017, pp. 23–30.
152. E. O. Neftci and B. B. Averbeck, "Reinforcement learning in artificial and biological systems," *Nature Machine Intelligence*, vol. 1, pp. 133–143, 2019.
153. D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, pp. 484–489, 2016.
154. O. Bousquet and A. Elisseeff, "Stability and generalization," *Journal of Machine Learning Research*, vol. 2, pp. 499–526, 2002.

155. R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Proc. 14th Int. Joint Conf. Artificial Intelligence*, 1995, pp. 1137–1143.
156. G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2021.
157. N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*. Cambridge, UK: Cambridge University Press, 2011.
158. D. H. Wolpert, “Stacked generalization,” *Neural Networks*, vol. 5, no. 2, pp. 241–259, 1992.
159. L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, 2nd ed. Hoboken, NJ, USA: Wiley, 2014.
160. J. Brownlee, *Imbalanced Classification with Python*. Vermont, Australia: Machine Learning Mastery, 2020.
161. Defence Research and Development Organisation, *Defence Science Journal: Instructions to Authors*. Delhi, India: DESIDOC, DRDO.