IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 28 February 2026, accepted 23 March 2026, date of publication 1 April 2026, date of current version 6 April 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3679745

RESEARCH ARTICLE

DyWIS: Dynamic Weighted Influence Score-Based Machine Learning for In-Game Win Loss Prediction in Elite Badminton Singles

NUR ATHIRAH MOHD SHAZALI[1], MOHD AMIRUDDIN ABD RAHMAN[1], (Member, IEEE), MUHAMMAD LUQMAN ARIF MOHAMAD[1], ATHIYYAH QISTINA ABDUL HAKIM[1], CACEJA ELYCA ANAK BUNDAK[1], SITI FAIRUZ MAT RADZI[1], KHAIRUL ADIB YUSOF[1], AND ROMI FADILLAH RAHMAT[2], (Member, IEEE)

[1]Department of Physics, Faculty of Science, Universiti Putra Malaysia (UPM), Serdang 43400, Malaysia
[2]Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan 20155, Indonesia

Corresponding authors: Mohd Amiruddin Abd Rahman (mohdamir@upm.edu.my) and Romi Fadillah Rahmat (romi.fadillah@usu.ac.id)

This work was supported by Universiti Putra Malaysia through the Geran Putra Berimpak under Grant GPB/2025/9817500.

ABSTRACT Badminton is a sport in which agility, anticipation, and tactical decision-making strongly influence competitive outcomes. Although recent work has shown that badminton performance and outcomes can be quantitatively modeled at the tournament and ranking levels, there remains limited research on rally-level modeling of in-game dynamics and real-time win probability prediction. Most existing studies rely on video-derived or biomechanical features, which are computationally expensive and unsuitable for real-time applications. This study addresses these gaps by introducing a interpretable and data-driven framework that supports lightweight real-time deployment that leverages rally-level score-difference dynamics to predict in-game outcomes for elite men's singles matches. A large-scale dataset comprising 18,742 games from the Badminton World Federation (BWF) tournaments between 2018 and 2023 was collected and refined to 1,021 games involving the top 10 ranked players. We propose a novel Dynamic Weighted Influence Score (DyWIS) framework that integrates machine learning classifiers with player-specific Absolute Win (AW) and Absolute Loss (AL) thresholds, and dynamically adjusts prediction confidence using historical win rates and current rally score progression. Experiments evaluated prediction performance at multiple rally checkpoints ($r = 20$ to 40), demonstrating that DyWIS provides informative early predictions with mean accuracy of above 84% at $r = 20$ for several top-performing classifiers, and approaches near-deterministic accuracy at later stages as score differences enter decisive AW/AL regions. At $r = 40$, DyWIS-KNN and DyWIS-linSVM achieved the highest mean accuracies of 93.37% and 93.19%, respectively, followed by DyWIS-XGBoost at 92.85%, consistently outperforming reimplemented baselines from prior work. Overall, the results indicate that DyWIS enables earlier and more stable win/loss detection as rallies progress, supporting real-time analytics, coaching decision support, and tactical visualization in professional badminton and other fast-paced individual sports.

INDEX TERMS Badminton, machine learning, classification, real-time prediction, sport analytics.

I. INTRODUCTION

Sport analytics [1] has emerged as a crucial tool for athletes, coaches, and teams to stay competitive in their respective fields. By transforming raw performance data into actionable insights, analytics supports diverse applications such as analyzing athlete performance [2], predicting game results [3], evaluating opponents' strengths and weaknesses [4], and improving team performance [5]. Machine learning (ML) methods in particular have proven effective for uncovering hidden patterns in complex sports data, enabling data-driven decision-making in both team and individual contexts [6], [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Ayman El-Baz.

A prominent use of sports analytics involves athlete profiling, which entails analyzing a single athlete's performances to gain insights into their typical performance in their respective field [8], [9], [10]. The findings can influence the performance of an individual athlete, the entire team for a particular game, or the entire tournament season [11]. Additionally, it can assist clubs in forecasting, assessing a player's market worth, and evaluating their risk of injury [12]. By providing these insights, sports analytics is transforming the way decisions are made in the world of sports.

The growing availability of structured and unstructured sports data has accelerated the adoption of predictive modeling across disciplines, including soccer [4], [7], [13], [14], [15], tennis [16], baseball [17], basketball [18], [19], cricket [1], [20] and handball [3], [21]. Within racket sports, analytics plays a vital role in capturing fast-paced dynamics where each rally provides a wealth of information about player strategy, momentum, and performance trends [22], [23].

Badminton, one of the fastest racket sports, with shuttlecock speeds exceeding 100 m/s [24], offers a unique challenge for predictive modeling. Since its inclusion in the 1992 Olympic Games, the sport has grown globally in both participation and competitive standards [25]. Its combination of speed, agility, and tactical variation makes outcome prediction particularly challenging [26]. Each rally can shift momentum, and subtle score differences often determine whether a player can maintain dominance or recover from a deficit [27]. As every rally and selected court area contains a wealth of information, leveraging this extensive collection of data to predict the key moments of the game is essential not only for improved viewing experiences but also for athlete training and strategic preparation [28]. Recent studies have explored video-based features such as strokes [29], [30], serves [31], trajectories [32], match statistics [33] and movement patterns [34], [35]. There are also studies on quantitative badminton performance using tournament and ranking levels [54], participation patterns [52], and scheduling [53] using Badminton World Federation (BWF) data. However, a real-time quantitative predictive approach using score-difference dynamics have not yet been studied for rally-level in-game outcomes. This represents a critical gap, as rally-level score progression directly reflects both player performance and psychological resilience under pressure.

Despite the rapid development of predictive analytics in sports [36], [37], several practical and methodological challenges remain in translating these models to real-time match environments. First, many existing systems rely on complex video processing pipelines that require high computational resources and manual data labeling, which limits their use during live tournaments [38]. Second, conventional ML approaches often lack contextual adaptability, treating all rallies as independent events without considering the evolving score momentum and player-specific tendencies. Third, the interpretability of closed-box models such as deep neural networks remains limited, making it difficult for coaches and analysts to understand or trust the underlying rationale behind predictions [39], [51]. Consequently, there is a growing demand for models that combine the predictive power of machine learning with the transparency of domain-informed rules, allowing clearer interpretation of decision boundaries during gameplay [40], [41]. Such models should be lightweight, interpretable, and context-aware, capable of integrating real-time score progression and historical player performance within a unified analytical framework. In this context, lightweight refers to the use of score-based inputs and simple decision mechanisms within the DyWIS framework, rather than restricting the choice of the underlying predictive model.

To address this gap, we propose a hybrid approach that integrates machine learning with Dynamic Weighted Influence Score (DyWIS), a novel framework designed to enhance predictive accuracy by incorporating both historical win rates and real-time score differences. In addition, we introduce player-specific absolute win (AW) and absolute loss (AL) thresholds, which enable early and reliable identification of decisive match scenarios. By focusing on the top 10 ranked men's singles players in the Badminton World Federation (BWF) tournaments, this study contributes a rigorous, data-driven framework for in-game win probability prediction. The key contributions of this study are as follows:

- We develop the DyWIS, which integrates machine learning classifiers with adaptive weighting based on historical win rates and current match dynamics.
- We analyze rally-by-rally score differences to identify thresholds that characterize AW and AL conditions for individual players.
- We incorporate player-specific AW/AL thresholds into DyWIS to enhance early detection of match outcomes and improve prediction stability.
- We demonstrate the effectiveness of the proposed models in generating real-time win probability predictions during rallies, outperforming conventional machine learning methods.

II. RELATED WORKS

Research in sports analytics has increasingly focused on classification and predictive modeling, with ML and deep learning (DL) emerging as powerful tools to forecast match outcomes, analyze performance, and inform strategic decisions. This section reviews relevant studies across competitive sports, with particular emphasis on badminton.

A. CLASSIFICATION IN COMPETITIVE SPORTS

Extensive research across multiple disciplines, including soccer, tennis, handball, baseball, and many more, has developed varied approaches to identifying predictive factors and improving model accuracy [42].

Predictive modeling in soccer has been extensively studied due to the availability of rich statistical data. Wong et al. [13] applied an extensive suite of machine learning models, including Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Convolutional Neural Networks (CNN) with ensemble methods using voting and stacking to classify outcomes in English Premier League (EPL) soccer matches. This work incorporated comprehensive features such as match statistics, fatigue levels, momentum indicators, and even weather conditions, aiming at binary classification of home team performance. Their CNN approach demonstrated the highest predictive accuracy at 63.7%. Notably, the SVM demonstrated balanced performance with an F1 score of 60.1%, which reflects its robustness in binary classification scenarios.

Chen et. al. [14] investigated multiclass prediction models for soccer outcomes using extensive engineered features, including FIFA ratings, historical statistics, Elo ratings, and betting odds from Europe's top five leagues. Among various tested ML models, the RF model emerged superior by achieving an accuracy of 56.25%. This underlines the utility of ensemble learning approaches, particularly when integrating diverse types of input features, ranging from rating systems to odds-based predictors.

Arntzen and Hvattum [15] presented an ordered logit regression combined with competing risk models to predict outcomes in football matches using Elo ratings and individual player plus-minus ratings within English leagues. Their optimal models achieved notably low information and quadratic loss values, which are 1.4721 and 0.6121, respectively, and closely align with bookmaker predictions, demonstrating robust statistical modelling performance that is competitive with industry-standard betting frameworks.

In handball, [43] adopted GBM to classify the match outcomes into three classes, such as home win, draw, or away win, using player strength ratings acquired from attack and defense, team form, and venue characteristics. Their model demonstrated an accuracy of approximately 73%, surpassing betting odds' predictive performance, which is around 69%. This demonstrates ML's capacity to effectively leverage structured, numeric datasets to surpass expert-based predictions.

Tennis has been a fertile testbed for classification methods because of its point-based structure [44], [45]. Gao and Kowalczyk [16] applied LR, SVM, and RF in tennis matches to utilize diverse variables, including serve accuracy, player's age, height, and playing surface. Notably, the RF classifier outperformed traditional bookmakers with an accuracy of 83.2%, considerably above the odds-based accuracy of 69%. The results highlighted the importance of technical match-level indicators. Similarly, studies integrating biomechanical or serve-tracking data have reported higher predictive accuracy than bookmakers, suggesting that ML can reliably capture nonlinear skill-performance relationships.

In baseball, Huang and Li [17] effectively employed artificial neural networks (ANN), one-dimensional CNNs (1DCNN), and SVM to classify outcomes of the game using a dataset derived from Baseball-Reference. This included detailed information on batting, pitching, and team home advantage. ANN emerged slightly superior with an accuracy of 94.18% compared to other techniques. This showcases the potential of neural networks in effectively managing and classifying large, feature-rich sports datasets.

In addition, ANNs, ensemble models, and hybrid systems have been successfully applied to basketball [46], cricket [1], and esports [47], each confirming that robust ML methods can surpass conventional heuristics. Collectively, these studies underscore that classification methodologies across diverse sports generally benefit from ensemble-based classifiers and neural networks algorithms, particularly effective when handling complex, high-dimensional, or nonlinear sports data. Ensemble techniques such as Random Forest and XGBoost consistently outperform simpler linear models due to their ability to capture nonlinear relationships and manage feature interactions effectively [48], [49]. Additionally, algorithms like ANN demonstrate strong performance in structured prediction tasks, especially when tuned to the characteristics of the dataset [50].

B. PREDICTIVE MODELING IN BADMINTON

Compared to major team sports, badminton analytics remains relatively underdeveloped. This creates additional effort for researchers who must extract the data themselves. Even so, research has been growing, particularly in predicting match outcomes, analyzing player actions, and anticipating future game events based on technical and tactical features. In addition to match-level predictive studies, several works have demonstrated that badminton performance and competitive outcomes can be quantitatively modeled using large-scale BWF data at the tournament and seasonal levels. Prior research has examined player rankings and ranking progression [51], tournament participation patterns [52], match scheduling effects [53], and performance efficiency across Olympic and non-Olympic cycles [54], [55]. These studies establish the feasibility of quantitative badminton analytics for understanding long-term competitive dynamics, although they do not address real-time, rally-level win probability prediction during live matches.

Sharma et. al. [33] analyzed badminton matches from prominent international tournaments, including the Australian, Malaysian, German, and Singapore Opens. Their study deployed innovative machine learning algorithms such as the Naive Bayes Classifier with Feature Weighting (NB-CBFW), CHIRP, and Hyper Pipes methods. The research uniquely included extensive match features, including match stages, player head-to-head statistics, duration, and comprehensive point, set, and game metrics. Among these, NB-CBFW demonstrated remarkably high predictive accuracy, achieving near-perfect performance in some scenarios and exhibiting a notably low root mean square error (RMSE) range from 0.006 to 0.10. Conversely, alternative models such as CHIRP and Hyper Pipes exhibited significantly lower accuracy and higher RMSE, thus emphasizing the efficacy of sophisticated Bayesian models in this context. This demonstrated the importance of carefully engineered features tailored to badminton match dynamics.

Yuan et al. [35] examined rally-level modeling in women's singles matches using match video data featuring an elite female player, An Se-young against the top five BWF-ranked players and match statistic from badminton competitions. The study applied a range of machine learning models including SVM with RBF kernel, KNN, Decision Tree, RF, and XGBoost to classify each rally segment as scoring or losing. Among them, the SVM with RBF kernel achieved the best accuracy at 87.5%, while KNN reached 75%. This study highlights the potential of video-derived rally-level features in accurately predicting performance outcomes, particularly when using refined algorithms like SVM, and reinforces the importance of model selection in maximizing predictive power in fast-paced match settings. In addition, recurrent neural networks (RNN) and related sequence models have been widely used in sports analytics to capture temporal dependencies in event sequences, particularly in settings where long observation windows and rich temporal features are available [56], [57].

Sheng et. al. [34] further analyzed match videos from international badminton competitions, focusing specifically on singles players among the world's top 10 male and female athletes. They employed a RF algorithm and leveraging detailed frequencies of 23 distinct technical actions, such as net-front plays, drop shots, pushes, and smashes. The analysis distinctly separated models by gender, reporting significant predictive accuracies where for male players, the RF model demonstrated an exceptional area under the curve (AUC) of 0.9726 with 87% accuracy, 93% sensitivity, and 80% specificity. In contrast, the female players' model exhibited a slightly lower AUC of 0.8730, accuracy of 83%, high sensitivity of 92%, but lower specificity of 68%. This gender differentiated performance underlines the importance of tailored approaches in predictive modeling to account for different play styles, physiological conditions, and strategic differences between male and female athletes.

Nokihara et. al. [2] focused on predicting future shuttlecock trajectories using Long Short-Term Memory (LSTM) combined with pose estimation techniques. Their work capitalized on match video-derived features, including shuttlecock position, player position, and seventeen distinct posture key points. The model significantly improved trajectory prediction accuracy and emphasizes the importance of integrating athlete positional and movement data for accurate, anticipatory analytics in fast-paced sports like badminton.

Ibh et al. [29] implemented a transformer encoder-decoder model to classify the next stroke type based on sophisticated player skeleton data and stroke embeddings. The model consistently outperformed conventional baselines and exhibited logical prediction patterns which achieved strong predictive performance with accuracy of 92.5% on the ShuttleSet dataset and 93.1% on the BadminDB dataset. These approaches highlight the potential of DL architectures for modeling temporal dependencies in high-speed rallies.

Some other works have also explored integrating tactical context. Studies have modeled rally segments as scoring vs losing [35], while others have examined serving patterns and player-specific strategies [31], [41]. Recent research on predictive modeling in badminton has predominantly focused on extracting and analyzing video-based features such as player movements [58], stroke types [59], and shuttlecock trajectories [60]. These works emphasize the interplay of technical, tactical, and situational factors in determining match outcomes.

While some research has incorporated statistical match data such as point and set metrics, there remains a noticeable gap in studies that focus specifically on score progression, point differentials, and rally-level scoring dynamics. For example, the lack of comprehensive datasets that combine player statistics, match conditions, and in-game events limits the ability to create predictive models tailored to badminton [61]. This gap highlights the need for targeted research that focuses on collecting and analyzing badminton-specific data to develop accurate win probability models. Future studies could benefit from exploring predictive models grounded in match statistics, which may offer valuable insights into performance trends that are not fully captured through video-based analysis alone.

C. COMPARATIVE ANALYSIS OF ML TECHNIQUES IN BADMINTON AND OTHER SPORTS

Table 1 provides a comparative summary of representative ML and DL models applied in sports prediction. Ensemble methods such as RF and XGBoost frequently outperform linear models due to their ability to capture nonlinear interactions and heterogeneous features, as seen in soccer [13], handball [43], and tennis [16]. In badminton, both handcrafted statistical features (e.g., NB-CBFW [33]) and video-derived technical actions (e.g., RF [34]) have achieved high predictive accuracy, while more recent studies using SVM and transformer architectures [35] demonstrate the potential of advanced models for rally-level prediction.

As summarized in Table 1, prior research demonstrates that ensemble models and neural networks often outperform simpler classifiers across a variety of sports. In badminton specifically, both handcrafted statistical features and video-based features have proven effective, but they typically depend on computationally expensive or dataset-specific inputs. This limits scalability and real-time deployment.

Despite promising advances, existing research in badminton analytics remains limited in its use of score-based features. Studies that are based on video-derived inputs such as strokes, trajectories, or skeletal data, which although effective, are computationally intensive, dataset-dependent, and difficult to scale for real-time applications. In contrast, rally-level score-difference dynamics are lightweight, universally available from official match records, and directly linked to both player dominance and psychological momentum.

TABLE 1. Comparison of ML models and results in sports reported in sports prediction literature.

Model	Sport Type	Features	Target	Accuracy/Result
Random Forest [34]	Badminton	23 technical actions such as Net Front, Slice/Drop, Push, Smash	Match outcome (binary)	Men's singles: • AUC = 0.9726 • Accuracy = 87% Woman's singles: • AUC = 0.8730 • Accuracy = 83%
NB-CBFW, CHIRP, Hyper Pipes [33]	Badminton	34 features including date, match stage, match duration, player stats, points, sets, H2H	Match outcome (binary)	Men's singles: • NB-CBFW = 100% • CHIRP = 90.47% • Hyper pipes = 97.64%
LR, Random Forest, SVM, XGB, GBM, CNN [13]	Soccer	52 features regarding team statistics, weather data, fatigue, momentum indicators	Home team match outcome (binary)	LR = 60.9% RF = 59.7% SVM = 62.9% XGB = 60.3% GBM = 61.4% CNN = 63.7%
Random Forest, XGB, CatBoost, Neural Network [43]	Handball	Team's strength (attack/defense), team form, venue, day until final	Home win, Draw, Away win (multi-class)	RF = 81.32% XGB = 73.57% CatBoost = 79.57% Neural Network = 68.08%
LR, SVM, Random Forest [16]	Tennis	Serve accuracy, age, height, surface, court condition, fatigue	Match outcome (binary)	LR = 62.06% SVM = 61.60% RF = 83.18%
SVM, KNN, DT, Random Forest, XGB [35]	Badminton	Number of shots, shot locations, techniques used in last 3 shots, scoring/loss position	Scoring result in each rally segment	SVM (RBF) = 87.5% KNN = 75% RF = 62.5% DT = 62.5% XGB = 56.25%

Surprisingly, these features have been underexplored for in-game prediction.

To address this gap, the present study introduces the DyWIS, a hybrid framework that integrates machine learning classifiers with player-specific absolute win/loss thresholds. By leveraging score-difference progression together with historical win rates, DyWIS offers an interpretable, data-efficient, and real-time-ready approach to win probability prediction in elite badminton.

III. METHODOLOGY

A. DATA AND PREPROCESSING

To develop and validate the proposed framework, we constructed a dataset from the official Badminton World Federation (BWF) website, which focuses on matches from the HSBC BWF World Tour series between 2018 and 2023 [62]. The raw dataset initially contained 18,742 games across all five competition categories (men's singles, women's singles, men's doubles, women's doubles, and mixed doubles).

The raw data extracted from the BWF website included several features, such as the tournament_name, country, date, duration, player1_name, player2_name, player1_nationality, player2_nationality, score progression in rally, and final match score. We started by applying a filtering process to the raw dataset to remove any matches where a player either retired or gave a walkover game, which typically occurs due to injuries or other unforeseen circumstances. These incomplete matches were excluded to ensure consistency in performance analysis.

Further, we restricted our analysis to the men's singles category. The scope is limited to matches involving the top 10 players in the HSBC Race to Finals ranking (week 48, 2023). This targeted selection provided a high-quality dataset of 1,021 games covering 121 players. The decision to focus on elite-level players ensures that the dataset reflects high-intensity matches and minimizes variability due to amateur or lower-tier performance.

Each match record included tournament information (name, country, and date), duration, player identities and nationalities, score progression at the rally level, and final outcomes. From these, we constructed a key feature called as the score difference, which represents the point margin between two players at any given rally, defined as

$$x_r = \text{player score at rally } r - \text{opponent score at rally } r \quad (1)$$

where x_r represents the score difference at rally r. In the context of badminton, a rally refers to a continuous sequence of play that begins with a serve and ends when a player wins the point, which either by grounding the shuttlecock within the opponent's court or forcing an error. As shown in Figure 1 where a rally is played by Player 1 and 2, and the rally is won by Player 1 after 5 shots, and Player 1 gets a point. In fact, the total number of rallies played will equal to the total number of score difference data in the game. Figure 2(a) shows the total number of rallies completed is 54 (with 28-26 as the final score), and equals to 54 score difference data as in Figure 2(b).

We recorded the progression of the score differences as the rally progresses in a game and constructed the score differences as a vector $X = \{x_1, x_2, \ldots, x_r\}$. This information

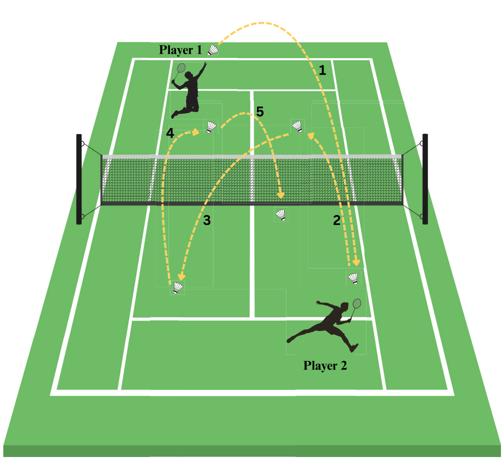


FIGURE 1. Example on how a rally in badminton is played. A rally is won by Player 1 after 5 shots.

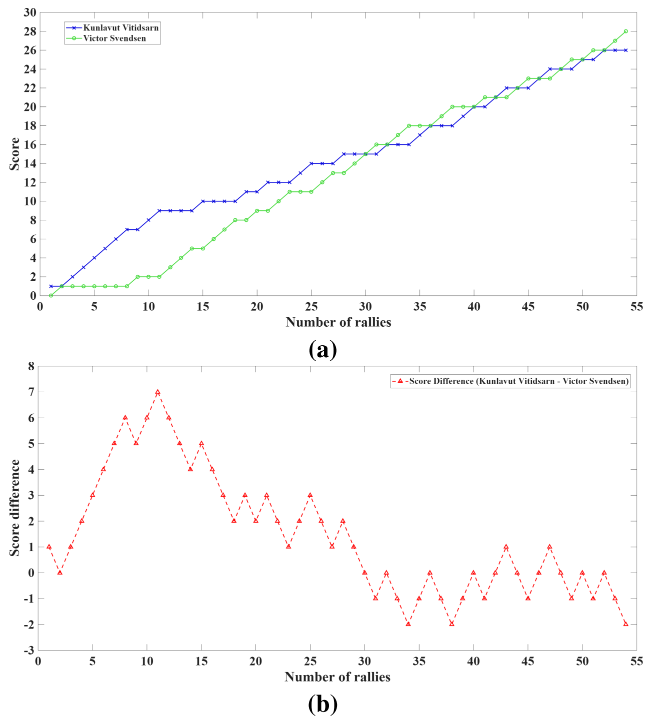


FIGURE 2. Example of score type between two players throughout a game. (a) Score progression between Kunlavut Vitidsarn and Victor Svendsen (b) Score differences of Kunlavut Vitidsarn over Victor Svendsen.

would provide a detailed view of how advantage shifts between players over time and forms the foundation for modelling in-game dynamics. Importantly, score difference does not encode tactical details such as stroke type, shot placement, court positioning, or biomechanical execution. Instead, it aggregates the downstream effects of these factors together with tactical pacing, error tolerance, psychological pressure management, and fatigue–recovery processes into a single, observable performance signal at each rally. We therefore use score difference as a lightweight proxy for competitive dynamics, suitable for real-time deployment. The score difference variations in a game are further explained

FIGURE 3. Number of rallies played in the dataset compared to the minimum and the maximum rallies possible in badminton.

TABLE 2. Key statistics of the processed dataset.

Metric	Value
Total matches	424
Total games	1021
Two-set games	251
Three-set games	173
Total rallies	44252
Minimum rally played	25
Minimum score point	21-4
Maximum rally played	54
Maximum score point	26-28
Average played rally	36.34
Shortest match	16 minutes
Longest match	113 minutes
Average match duration	58 minutes
Minimum score difference	−16
Maximum score difference	18
Average score difference	−1.67
Dataset timeframe	2018–2023
Number of tournaments	82
Number of player's countries	20

as in Figure 2(b) which shows how the value of the score difference changes when there is score progression (Figure 2(a)) between two players. This representation provides a lightweight proxy for momentum shifts and competitive dynamics. The final positive score difference indicates that the player has won the game, while a negative one indicates a loss.

A comprehensive summary of the dataset used in this study is presented in Table 2. Statistical analysis highlights the minimum, maximum, and average values of rally count, score difference, and match duration, providing a clear overview of the competitive dynamics represented in the data. At any point in a game, the number of rallies played or number of score diffrences can be computed by summing the current scores of both players. For example, a final score of 21–4 corresponds to 25 rallies, whereas a tightly contested score of 26–28 indicates 54 rallies. The dataset encompasses a wide range of match intensities, as illustrated in Figure 3, where the distribution of rally counts is compared with the theoretical minimum (21–0, 21 rallies) and maximum (30–29, 59 rallies) possible in badminton scoring. Across all matches, score differences ranged from –18 to +15, with an average of –1.67, indicating a slightly higher frequency of negative margins among players. While small positive or negative score differences may occur frequently during rallies, as shown in Figure 4, larger absolute differences are typically decisive, signaling winning or losing the game. This score-difference indicator serves as a key parameter in the proposed DyWIS framework, forming the foundation for distinguishing between winning and losing trends within a match.

TABLE 3. Dataset description.

Player	Country	HSBC Rank (2023)	Game Margin	Score	Score Difference	Number of Games	Total Games Win	Total Games Loss	Total Rally Win	Total Rally Loss	Train Data Win	Train Data Loss	Train Data (After SMOTE) Win	Train Data (After SMOTE) Loss	Test Data	AW	AL
Kodai Naraoka	Japan	1	Largest win	21–6	15	128	75	53	2675	1924	68	22	68	68	38	8	-10
			Smallest win	22–20	2												
			Largest loss	6–21	-15												
			Smallest loss	19–21	-2												
Jonatan Christie	Indonesia	2	Largest win	21–6	15	127	73	54	2634	2011	64	25	64	64	38	8	-10
			Smallest win	24–22	2												
			Largest loss	9–21	-12												
			Smallest loss	19–21	-2												
Shi Yu Qi	China	3	Largest win	21–5	16	99	64	35	2320	1309	58	11	58	58	30	6	-8
			Smallest win	21–19	2												
			Largest loss	5–21	-16												
			Smallest loss	20–22	-2												
Li Shi Feng	China	4	Largest win	21–3	18	137	77	60	2765	2229	69	27	69	69	41	8	-6
			Smallest win	24–22	2												
			Largest loss	9–21	-12												
			Smallest loss	19–21	-2												
Viktor Axelsen	Denmark	5	Largest win	21–6	15	40	28	12	979	474	25	3	25	25	12	5	-11
			Smallest win	21–19	2												
			Largest loss	10–21	-11												
			Smallest loss	19–21	-2												
Anthony S. Ginting	Indonesia	6	Largest win	21–5	16	95	61	34	2209	1276	55	12	55	55	29	10	-7
			Smallest win	21–19	2												
			Largest loss	8–21	-13												
			Smallest loss	22–24	-2												
Anders Antonsen	Denmark	7	Largest win	21–7	14	160	88	72	3153	2556	79	33	79	79	48	10	-8
			Smallest win	25–23	2												
			Largest loss	6–21	-15												
			Smallest loss	20–22	-2												
Kenta Nishimoto	Japan	8	Largest win	21–8	13	231	117	114	4274	4201	106	56	106	106	69	8	-8
			Smallest win	21–19	2												
			Largest loss	6–21	-15												
			Smallest loss	19–21	-2												
Kunlavut Vitidsarn	Thailand	9	Largest win	21–3	18	122	64	58	1424	2112	57	28	57	57	37	8	-6
			Smallest win	21–19	2												
			Largest loss	7–21	-14												
			Smallest loss	19–21	-2												
Ng Tze Yong	Malaysia	10	Largest win	21–4	17	105	63	42	2215	1512	58	16	58	58	31	10	-10
			Smallest win	21–19	2												
			Largest loss	7–21	-14												
			Smallest loss	20–22	-2												

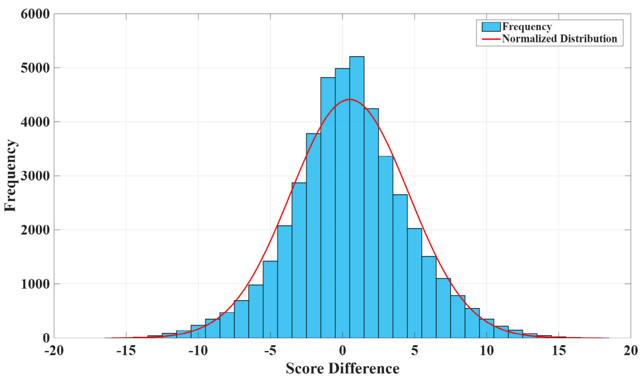


FIGURE 4. Frequency distribution of score differences from the dataset.

B. DETERMINATION OF ABSOLUTE WIN (AW) AND ABSOLUTE LOSS (AL)

It is possible to determine the confirmed winning or losing score by investigating score difference pattern in each player's game history. First, we analyzed the score difference of each player's game to determine the highest and lowest score difference and stored as $g_m = (max(x_r), min(x_r))$, where g_m is a game of index m. We named $max(x_r)$ and $min(x_r)$ as the highest lead and lag score in game m respectively. Then, we sorted the g_m in ascending order (more positives) of $max(x_r)$ for each losing games and continued with descending order of winning games. We observed the $max(x_r)$ value in the last indexed g_m in losing games to determine confirmation of a player would win a game. To visualize this, we plotted g_m as in Figure 5(a). The figure uses Kenta Nishimoto games as an example. We could confirm that Nishimoto would win a game whenever he obtained the highest lead score difference of 8 as it would be possible for him to still lose a game eventhough he already achieved lead score difference of 7. This means that when he leads by ≥ 8, victory was assured. We determined this as Absolute Win (AW) value.

Similarly, to determine the decisive losing score difference value, we use similar set of data and rearrganged it as $g_n = (min(x_r), max(x_r))$, where n is the new game index. Then, we sorted the g_n in descending order (more negatives) of $min(x_r)$ for each losing games and continued with ascending order of winning games. We observed the $min(x_r)$ value in the first indexed g_n in winning games to determine confirmation of a player would lose a game. This case is illustrated in Figure 5(b). As could be seen at the value of -7, Nishimoto could still win a game, but when he obtained the lag score of -8 he would never win. Thus, -8 is determined as the Absolute Loss (AL) value for him. The value of AW and AL for every players are recorded as in Table 3. To standardize the notation with the score difference, we equate AW and AL as x_w and x_l respectively.

The resulting AW and AL values are player-conditioned thresholds and empirically derived from individual players' past competitive data. Concretely, we analyze the maximum and minimum score differences attained in each game and

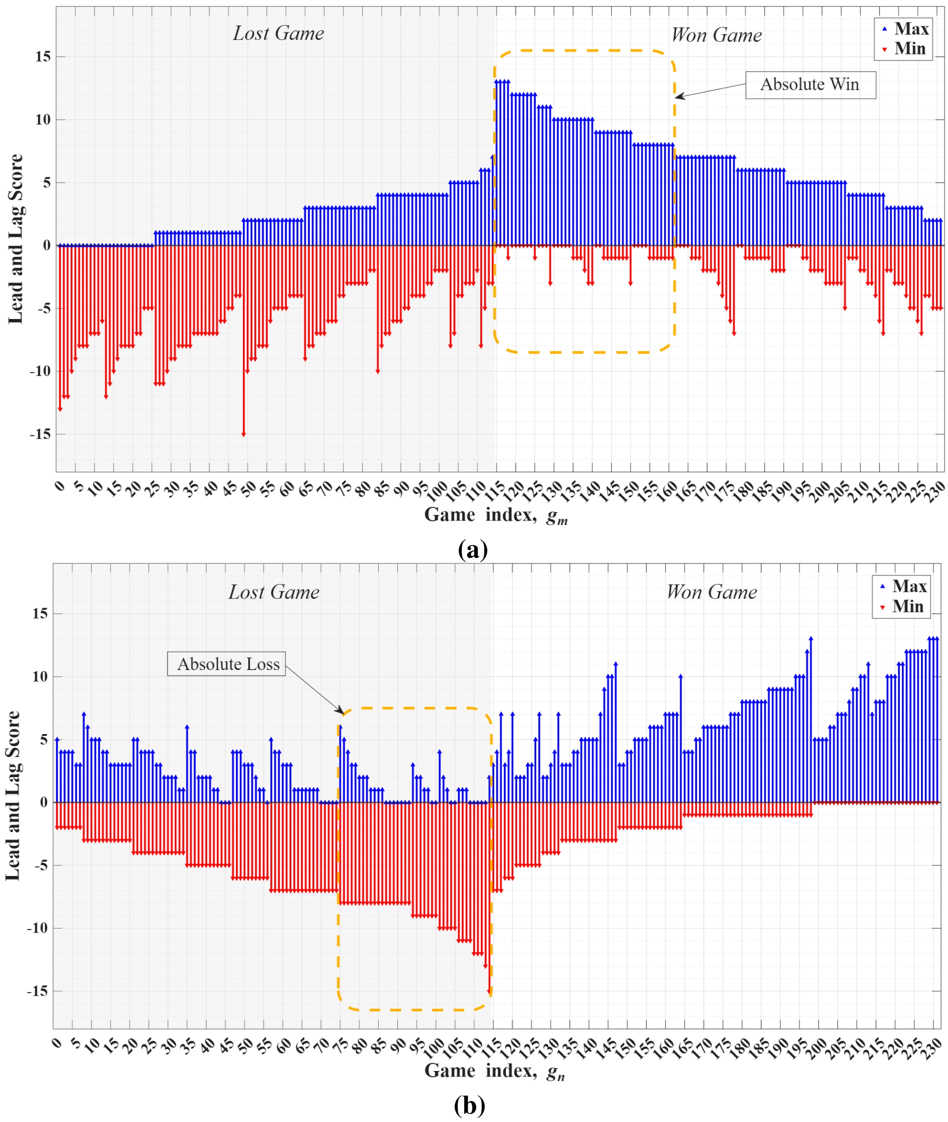


FIGURE 5. Arranged maximum lead and lag score differences in Kenta Nishimoto's matches to determine (a) Absolute Win (AW) and (b) Absolute Loss (AL) conditions.

identify the smallest positive lead and largest negative deficit beyond which no historical comebacks are observed. This player-conditioned strategy captures individual dominance and recovery characteristics without introducing free tuning hyperparameters, as the thresholds are directly derived from each player's historical data and remain fixed during inference.

C. DYNAMIC WEIGHTED INFLUENCE SCORE

The core contribution of this study is the introduction of the Dynamic Weighted Influence Score (DyWIS), a hybrid framework that integrates player-specific score influence probability with machine learning (ML) classifiers based on their historical score difference pattern. The objective is to improve the accuracy, stability, and interpretability of in-game win probability prediction in badminton.

Conventional machine learning (ML) models treat all rally states uniformly, often overlooking *momentum shifts* that provide early cues of victory or defeat. For instance, when a player leads by a large margin, the outcome becomes nearly deterministic, whereas narrow margins demand higher sensitivity to subtle performance fluctuations. The proposed DyWIS addresses this limitation by integrating four score-influenced parameters that jointly enhance prediction accuracy:

1) Absolute win and loss thresholds (x_w and x_l): Directly classify extreme score conditions as decisive outcomes.
2) Machine learning confidence score ($p(X_{tr})$): Represents the classifier-derived probability of winning based on the score-difference vector.
3) Historical win-rate weighting (w_r): Incorporates player-specific prior performance to adjust predictive confidence.

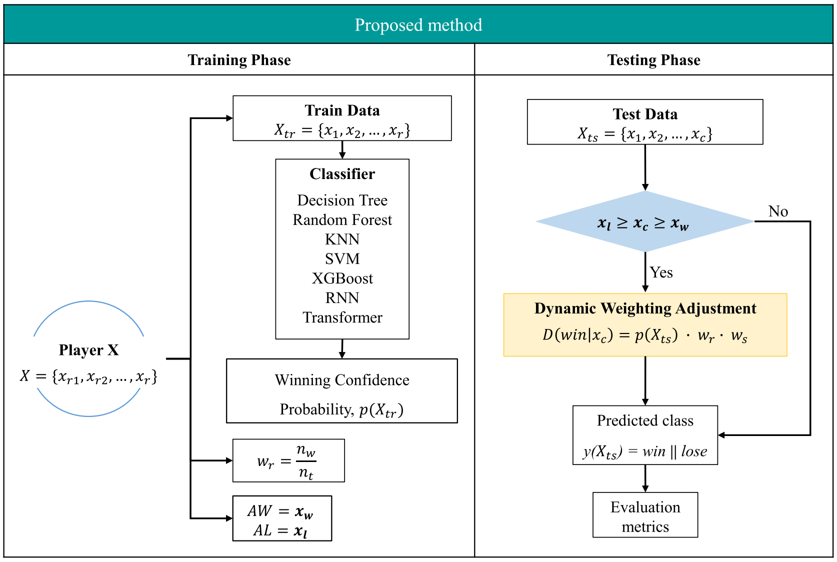


FIGURE 6. Proposed framework of DyWIS.

4) Current score influence weighting (w_s): Dynamically adjusts the confidence score according to the player's progress toward the minimum winning score.

The first three parameters can be computed before the match begins, while the fourth is calculated in real time as the game progresses. This hybrid formulation enables DyWIS to provide earlier, more stable, and context-aware outcome predictions, especially in ambiguous rally situations where conventional models typically struggle.

The score-difference vector X_{tr} for each player was used to train multiple machine learning and deep learning classifiers to estimate the winning confidence probability, denoted as $p(X_{tr})$. Seven classifiers were evaluated in this study: Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor (KNN), XGBoost (XGB), Support Vector Machine (SVM), Recurrent Neural Network (RNN) and Transformer-based model. The composition of each player's training dataset is summarized in Table 3. The table shows the players with their game statistics, which were analyzed in this study. Each player has a different number of matches, based on their individual participation history in the selected tournaments.

Preliminary analysis revealed a substantial imbalance in outcomes across players. For example, Kodai Naraoka recorded wins in 75.6% of matches, while only 24.4% were losses. Such imbalance risks biasing classifiers toward the majority class. To mitigate this, we employed the Synthetic Minority Over-sampling Technique (SMOTE) to generate synthetic training samples for the minority class, balancing win/loss distributions at a 50:50 ratio. This step improves generalization and ensures that models remain sensitive to both outcomes. Although rally-level score data are sequential, SMOTE is applied at the score-difference feature level to balance win–loss samples for the classifiers in accordance with the DyWIS formulation, which does not explicitly model rally-to-rally transitions.

TABLE 4. Model parameters and hyperparameters used in this study.

Classifier	Parameter	Value
Decision Tree	Criterion	Gini impurity
	Min leaf size	1
	Min parent size	10
	Max splits per tree	135
Random Forest	Number of trees	50
	Min leaf size	1
k-NN	Number of neighbors	10
SVM	Linear kernel	
	Box constraint (C)	1
	Polynomial kernel	
	Degree	3
	Coefficient	0
	Gamma	1
	Box constraint (C)	1
	RBF kernel	
	Gamma	1
	Box constraint (C)	1
XGBoost	Booster type	gbtree
	Learning rate	0.1
	Number of boosting rounds	100
	Maximum tree depth	5
RNN	Number of layers	single LSTM layer
	Hidden units	32
	Optimizer	Adam
	Learning rate	0.001
Transformer	Number of layers	1 encoder block
	Attention heads	4
	Embedding dimension (d_{model})	32
	Optimizer	Adam
	Learning rate	0.001

Parameters for each model were selected empirically through preliminary testing to balance accuracy and generalization as summarized in Table 4. A moderate learning rate ($\eta = 0.1$) and shallow tree depth ($d = 5$) were used for XGBoost to prevent overfitting, while the SVM box constraint ($C = 1$) provided a balanced trade-off between margin maximization and training accuracy. For the sequence-based RNN and Transformer models, simple baseline configurations were adopted to demonstrate

Algorithm 1 DyWIS Algorithm

Input: Score difference data X_{tr}
Output: Predicted game outcome $y(X_{ts})$
Extract feature matrix X and target label y from input data
Initialize rally set r
Initialize Absolute Win threshold x_w and Absolute Loss threshold x_l
Split data into training set $(X_{\text{tr}}, y_{\text{tr}})$ and testing set $(X_{\text{ts}}, y_{\text{ts}})$
Compute total games n_t, total winning games n_w, and historical win rate $w_r = n_w/n_t$
Balance the training data using SMOTE:
  Identify minority class samples in y_{tr}
  Generate synthetic samples by interpolating between random minority samples
  Combine and shuffle to form the balanced dataset $(X_{\text{tr\_balanced}}, y_{\text{tr\_balanced}})$
for each game g **do**
  for each rally r in the game **do**
    Compute score influence factor $w_s = min(\frac{score\ at\ rally\ r}{21}, 1)$
    if current score difference $x_c \geq x_w$ **then**
      Classify outcome $y(X_{ts}) = 1$
    else if $x_c \leq x_l$ **then**
      Classify outcome $y(X_{ts}) = 0$
    else
      Initialize model parameters according to Table 4
      Train model using balanced training set
      Predict win probability $p(X_{ts})$ for current rally r
      Compute adjusted win probability: $D(\text{win} \mid x_c) = p(X_{ts}) \cdot w_r \cdot w_s$
      if $D(\text{win} \mid x_c) \geq 0.5$ **then**
        Classify outcome $y(X_{ts}) = 1$
      else
        Classify outcome $y(X_{ts}) = 0$
      end if
    end if
  end for
  Store each game final outcome f
end for
Compute evaluation metrics: Accuracy, Precision, Recall, F1-score
return Predicted outcomes $y(X_{ts})$ and performance metrics

framework compatibility rather than to optimize deep model performance. Unless otherwise stated, all other parameters were kept at their default Python library values.

During training, the shortest match length (in rallies) within the dataset was used to standardized sequence length for all players. This ensured that every model learned from consistent feature dimensions while enabling early in-game prediction. By constraining the input to the minimum number of rallies observed, the framework can generate reliable outcome predictions as early as possible during gameplay.

We calculated the player's historical winning rate, w_r, to encode long-term performance given as

$$w_r = \frac{n_w}{n_t} \tag{2}$$

where n_w represents the number of matches won by the player and n_t is the total games played. For example, as in Table 3, Kodai won 75 out of 128 games, which translates to his w_r score of 0.5859.

During the game, a player will be assigned with a Dynamic Weighting Adjustment (DWA) value after each rally by manipulating the current score difference, x_c. If the x_c does not reach the x_w or x_l value, the DWA value could be calculated as the product of the classifier-derived confidence score and in-game contextual weightings as follows:

$$D(win|x_c) = p(X_{ts}) \cdot w_r \cdot w_s \tag{3}$$

where w_s is the score influence factor which infer the score nearer to winning score of the current game, normalized by the minimum winning score of 21 points. The w_s is capped at 1.0 once the player reaches 21 points. This prevents excessive weighting in deuce situations, where the score can extend beyond 21 (up to 30) but the game outcome is determined by the two-point lead rule, rather than by continued linear score accumulation. In practice, capping w_s at 1.0 for scores $\geq$21 does not materially affect the predictions because once the game enters deuce, the decisive transitions occur when one player establishes a two-point advantage, at which point the game terminates. This fixed normalization ensures that the influence of score progression remains interpretable and comparable across matches and players. A player would have a big w_s if his score nearer to the winning score. For example, a player that has reached a score of 13 in a game would have $w_s = 0.6190$ compared to when only reached a score of 5 ($w_s = 0.2381$) in a game. Using the official winning score as the normalization constant avoids introducing additional tuning parameters, preserving numerical stability and supporting real-time deployment without the need for match-specific calibration. Then, the winning or losing class could be determined as

$$y(X_{ts}) = \begin{cases} 1, & \text{if } x_c \geq x_w \text{ or } D(\text{win} \mid x_c) \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where 1 represents win class and 0 represents loss class.

The detailed implementation of the proposed DyWIS framework is outlined in Algorithm 1. During a match, DyWIS updates the win–loss confidence online by recomputing the dynamic weighting terms based on the current rally state, while the underlying classifier is trained offline using historical training games. Thus, real-time inference only requires evaluating the pre-trained model on the updated score-difference sequence and applying the DyWIS weighting and AW/AL decision rules without retraining the model at each rally. Figure 6 illustrates the proposed DyWIS framework.

IV. RESULTS AND DISCUSSION

Accurate early prediction is essential in live match environments, where the goal is not merely to identify the final winner but to provide stable and interpretable win probabilities as rallies unfold. Such real-time predictions enable coaches and analysts to detect shifts in game momentum and adjust strategies dynamically. To achieve this, the proposed DyWIS framework integrates dynamic weighting adjustments and AW and AL thresholds within conventional machine or deep learning models. This combination allows continuous updates of predicted outcomes after each rally, ensuring that the output reflects the current game flow and intensity rather than static match statistics. The framework is designed to enhance both predictive accuracy and interpretability while maintaining low computational cost that is a critical requirement for deployment in live sports analytics applications.

A. PREDICTION PERFORMANCE OF DYWIS AT DIFFERENT RALLY STAGES

We evaluated the framework on the processed men's singles dataset (2018–2023, 1021 games, 44252 rallies). In the context of live-match analytics, the critical question is how early a prediction can be considered reliable. Because DyWIS continuously updates win probabilities after each rally, it allows examination of how prediction accuracy evolves as the match progresses. To study this, we evaluated all classifiers at five rally checkpoints, $r = \{20, 25, 30, 35, 40\}$. This range was selected based on the dataset's average rally length of 36.34 and its observed span from 25 to 54 rallies per match (Table 2). The inclusion of the 20th rally checkpoint was intentional to assess model performance in early-game scenarios, where accurate predictions are most valuable for real-time applications.

The accuracy and F1-score for every player are listed in Table 5. Reporting both metrics provides a more balanced evaluation, and this allows us to determine whether the model maintains fair discrimination between win and loss outcomes or simply favors the majority class. In particular, a scenario where accuracy remains high but the F1-score declines suggests bias toward predicting wins, whereas simultaneous improvement in both metrics indicates genuine model learning and balanced classification. For clarity, the strongest performance results for mean accuracy are highlighted in bold, indicating the top-performing outcomes across classifiers and rally stages for both accuracy and F1-score.

The first key observation is that DyWIS already effective as early as $r = 20$ rallies. At this point, most players achieve accuracies above 80% for several classifiers, with the main exceptions being the linSVM and polySVM variants. Among all models, DyWIS-RF, DyWIS-rbfSVM, DyWIS-XGBoost, DyWIS-RNN, and DyWIS-Transformer consistently deliver the strongest early-stage performance, with mean accuracies across players exceeding 84% (Table 5). In badminton, to obtain good prediction at 20 rallies are crucial since the minimum rally for an outcome is 21 (see Figure 3). Thus, if early predictions indicate an unfavorable outcome, coaches and players could adjust tactics before the game outcome becomes irreversible.

On top of that, the corresponding F1-scores remain high (typically $>$ 88%), confirming that the models effectively distinguish between both win and loss classes rather than favoring a single outcome. The ROC analysis in Figure 7 confirms this observation from a threshold-independent perspective, indicating that even at early game stages, DyWIS achieves reliable class separation. These results show that early predictions are already informative, and performance continues to improve as more rallies are incorporated. When comparing classical machine learning models with sequence-based approaches, DyWIS-RNN and DyWIS-Transformer exhibit performance that is competitive with, but not consistently superior to the strongest conventional classifiers at early rally stages. At $r = 20$, their accuracy and F1-score are generally comparable to DyWIS-RF and DyWIS-rbfSVM. These findings suggest that the benefit of modeling long-term temporal dependencies in this setting may be limited especially when only short rally sequences are available.

As the rally length increases to $r = 25$ and $r = 30$, the prediction performance improves further for nearly all players and classifiers. For instance, Kodai's DyWIS-KNN accuracy rises from 81.58% at $r = 20$ to 89.47% at $r = 25$ and remains consistently high at $r = 30$. Similarly, Jonatan's DyWIS-DT momentarily reaches 94.74% at $r = 30$. DyWIS-RNN shows gradual improvements in both accuracy and F1-score for several players, consistent with its ability to accumulate temporal evidence across rallies. In contrast, DyWIS-Transformer does not consistently outperform either RNN or the strongest tree-based classifiers, reflecting that the added architectural complexity does not necessarily translate into performance gains under the relatively short rally sequences typical of badminton matches. A comparable trend is observed for Li Shi Feng and Anthony S. Ginting, for whom $r = 30$ represents a turning point where the score difference becomes sufficiently distinctive for reliable outcome prediction. At this stage, even relatively simple classifiers such as DyWIS-DT, DyWIS-RF, and DyWIS-linSVM can effectively separate winning and losing margins. As the game progresses, the score influence term w_s amplifies the classifier confidence, reinforcing the accuracy and stability of predictions across longer rallies.

At $r = 35$ and $r = 40$, results approach near-perfect accuracy. It is important to interpret the very high accuracies observed at later rally checkpoints in the context of the scoring structure and the AW/AL thresholds. At these stages, a large proportion of rallies fall into regions where the current score difference, x_c already exceeds the player-specific AW/AL threshold. By construction, such states

TABLE 5. Classifier performance per player across rally, $r = 20, 25, 30, 35, 40$ for accuracy, and F1-score.

Dataset	Classifier	Accuracy (%)					F1-score (%)				
		20	25	30	35	40	20	25	30	35	40
Kodai	dt	81.58	86.84	86.84	84.21	92.11	89.86	92.54	92.06	89.66	94.92
	knn	81.58	89.47	89.47	92.11	97.37	89.86	93.94	93.94	95.38	98.41
	rf	81.58	84.21	84.21	84.21	84.21	89.86	91.18	91.18	91.18	91.18
	linSVM	89.47	89.47	92.11	92.11	97.37	93.94	93.75	95.24	95.38	98.41
	polySVM	78.95	78.95	76.32	73.68	86.84	87.10	87.10	85.25	83.87	92.31
	rbfSVM	81.58	84.21	84.21	84.21	84.21	89.86	91.18	91.18	91.18	91.18
	XGBoost	81.58	89.47	92.11	89.47	100.00	89.86	93.94	95.38	93.33	100.00
	RNN	81.58	86.84	89.47	81.58	92.11	89.86	92.31	93.55	88.14	95.24
	Transformer	81.58	84.21	84.21	84.21	84.21	89.86	91.18	91.18	91.18	91.18
Jonatan	dt	81.58	84.21	94.74	89.47	86.84	89.23	90.63	96.55	92.86	90.57
	knn	81.58	86.84	89.47	97.37	92.11	89.23	92.06	93.33	98.31	95.08
	rf	81.58	81.58	89.47	89.47	89.47	89.23	89.23	93.55	93.55	93.55
	linSVM	84.21	84.21	89.47	92.11	84.21	89.66	88.89	92.86	94.55	89.29
	polySVM	78.95	86.84	86.84	97.37	92.11	86.67	91.53	91.23	98.31	94.74
	rbfSVM	81.58	81.58	89.47	89.47	89.47	89.23	89.23	93.55	93.55	93.55
	XGBoost	81.58	86.84	89.47	92.11	84.21	89.23	92.06	92.86	94.74	88.89
	RNN	81.58	92.11	92.11	84.21	92.11	89.23	95.08	94.74	88.46	94.55
	Transformer	81.58	81.58	89.47	89.47	89.47	89.23	89.23	93.55	93.55	93.55
Shi Yu Qi	dt	90.00	83.33	76.67	90.00	93.33	94.12	88.89	84.44	93.62	95.65
	knn	90.00	90.00	86.67	83.33	86.67	94.12	93.88	91.67	88.89	91.30
	rf	90.00	90.00	90.00	90.00	90.00	94.12	94.12	94.12	94.12	94.12
	linSVM	76.67	76.67	76.67	80.00	86.67	85.11	84.44	84.44	86.96	91.30
	polySVM	76.67	83.33	83.33	86.67	86.67	85.11	89.80	89.36	91.67	91.67
	rbfSVM	90.00	90.00	90.00	90.00	90.00	94.12	94.12	94.12	94.12	94.12
	XGBoost	90.00	80.00	80.00	80.00	93.33	94.12	86.96	86.96	86.36	95.65
	RNN	90.00	86.67	83.33	86.67	93.33	94.12	92.00	89.36	91.30	95.65
	Transformer	90.00	90.00	90.00	83.33	90.00	94.12	94.12	94.12	89.80	93.62
Li Shi Feng	dt	82.93	82.93	87.80	73.17	85.37	90.41	90.14	92.31	82.54	90.63
	knn	82.93	78.05	87.80	90.24	87.80	90.41	87.32	92.96	94.29	92.31
	rf	82.93	82.93	85.37	87.80	87.80	90.41	90.41	91.67	92.96	92.96
	linSVM	80.49	80.49	78.05	82.93	92.68	86.67	87.10	85.25	88.89	95.52
	polySVM	80.49	80.49	80.49	82.93	82.93	88.24	88.24	87.88	89.55	89.55
	rbfSVM	82.93	82.93	85.37	87.80	87.80	90.41	90.41	91.67	92.96	92.96
	XGBoost	82.93	85.37	87.80	85.37	85.37	90.41	91.43	92.31	90.91	90.63
	RNN	82.93	82.93	82.93	78.05	90.24	90.41	90.14	88.89	85.25	93.75
	Transformer	82.93	82.93	85.37	87.80	87.80	90.41	90.41	91.67	92.96	92.96
Viktor	dt	100.00	66.67	83.33	83.33	83.33	100.00	71.43	87.50	87.50	87.50
	knn	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	rf	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	linSVM	58.33	58.33	75.00	66.67	100.00	61.54	61.54	80.00	71.43	100.00
	polySVM	91.67	91.67	91.67	75.00	100.00	94.12	94.12	94.12	80.00	100.00
	rbfSVM	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	XGBoost	100.00	83.33	91.67	91.67	100.00	100.00	87.50	94.12	94.12	100.00
	RNN	100.00	91.67	100.00	100.00	100.00	100.00	94.12	100.00	100.00	100.00
	Transformer	100.00	100.00	100.00	83.33	100.00	100.00	100.00	100.00	87.50	100.00
Anthony	dt	75.00	82.14	82.14	89.29	85.71	85.11	89.36	88.89	93.33	90.48
	knn	75.00	82.14	78.57	85.71	92.86	85.11	88.89	86.36	90.91	95.45
	rf	78.57	78.57	82.14	82.14	82.14	88.00	88.00	89.80	89.80	89.80
	linSVM	67.86	78.57	75.00	82.14	89.29	76.92	85.71	82.93	88.37	93.02
	polySVM	75.00	78.57	78.57	82.14	89.29	82.93	86.36	86.36	88.37	93.33
	rbfSVM	78.57	78.57	82.14	82.14	82.14	88.00	88.00	89.80	89.80	89.80
	XGBoost	78.57	82.14	82.14	85.71	89.29	87.50	88.89	88.37	90.91	93.02
	RNN	75.00	82.14	78.57	89.29	89.29	85.11	88.89	85.71	93.33	93.02
	Transformer	78.57	78.57	82.14	82.14	89.29	88.00	88.00	89.80	89.36	93.33
Anders	dt	81.25	83.33	91.67	75.00	95.83	89.66	90.70	95.12	82.86	97.37
	knn	81.25	83.33	91.67	93.75	95.83	89.66	90.70	95.12	96.30	97.50
	rf	81.25	81.25	81.25	83.33	83.33	89.66	89.66	89.66	90.70	90.70
	linSVM	77.08	87.50	89.58	93.75	97.92	86.08	92.50	93.67	96.20	98.73
	polySVM	79.17	70.83	79.17	87.50	91.67	87.50	82.05	87.18	92.50	94.87
	rbfSVM	81.25	81.25	81.25	83.33	83.33	89.66	89.66	89.66	90.70	90.70
	XGBoost	81.25	85.42	89.58	93.75	93.75	89.66	91.76	93.67	96.20	96.20
	RNN	81.25	83.33	85.42	79.17	89.58	89.66	90.70	90.91	86.11	93.33
	Transformer	81.25	81.25	81.25	83.33	83.33	89.66	89.66	89.66	90.70	90.70
Kenta	dt	85.51	88.41	89.86	85.51	91.30	92.06	93.55	94.02	90.74	94.55
	knn	85.51	88.41	86.96	92.75	95.65	92.06	93.55	92.44	95.73	97.44
	rf	85.51	88.41	89.86	89.86	89.86	92.06	93.55	94.31	94.31	94.31
	linSVM	82.61	88.41	89.86	89.86	95.65	90.00	93.10	93.91	93.81	97.39
	polySVM	79.71	86.96	88.41	88.41	91.30	87.72	92.44	93.22	93.10	94.83
	rbfSVM	85.51	88.41	89.86	89.86	89.86	92.06	93.55	94.31	94.31	94.31
	XGBoost	85.51	88.41	89.86	89.86	97.10	92.06	93.55	94.02	93.81	98.25
	RNN	85.51	88.41	88.41	88.41	91.30	92.06	93.55	93.22	92.86	94.64
	Transformer	85.51	88.41	89.86	89.86	89.86	92.06	93.55	94.31	94.31	94.31
Kunlavut	dt	81.08	81.08	83.78	81.08	86.49	89.55	89.55	90.00	87.72	90.91
	knn	81.08	83.78	86.49	89.19	91.89	89.55	90.91	92.06	93.55	95.08
	rf	81.08	81.08	83.78	86.49	86.49	89.55	89.55	90.91	92.31	92.31
	linSVM	81.08	81.08	86.49	94.59	94.59	88.52	87.72	92.06	96.67	96.55
	polySVM	86.49	91.89	91.89	94.59	97.30	91.80	95.08	95.08	96.67	98.36
	rbfSVM	81.08	81.08	83.78	86.49	86.49	89.55	89.55	90.91	92.31	92.31
	XGBoost	81.08	86.49	86.49	97.30	91.89	89.55	92.31	92.06	98.31	94.74
	RNN	81.08	86.49	83.78	89.19	91.89	89.55	92.31	90.32	93.33	94.74
	Transformer	81.08	81.08	83.78	86.49	86.49	89.55	89.55	90.91	92.31	92.31
Ng Tze Yong	dt	80.65	80.65	74.19	74.19	93.55	89.29	88.89	84.00	83.33	96.00
	knn	80.65	83.87	70.97	77.42	93.55	89.29	90.91	82.35	86.27	96.30
	rf	83.87	83.87	83.87	83.87	83.87	91.23	91.23	91.23	91.23	91.23
	linSVM	70.97	74.19	74.19	80.65	93.55	82.35	84.62	84.00	88.89	96.15
	polySVM	74.19	77.42	74.19	74.19	90.32	84.62	86.27	84.00	84.00	94.34
	rbfSVM	83.87	83.87	83.87	83.87	83.87	91.23	91.23	91.23	91.23	91.23
	XGBoost	80.65	80.65	74.19	77.42	93.55	89.29	88.89	83.33	85.71	96.00
	RNN	83.87	77.42	70.97	77.42	83.87	91.23	86.79	80.00	85.71	90.20
	Transformer	83.87	83.87	83.87	83.87	96.77	91.23	91.23	91.23	91.23	98.11
Mean	dt	83.96	81.96	85.10	82.53	89.39	90.93	88.57	90.49	88.42	92.86
	knn	83.96	86.59	86.81	90.19	**93.37**	90.93	92.22	92.02	93.96	**95.89**
	rf	84.64	85.19	87.00	87.72	87.72	91.41	91.69	92.64	93.01	93.01
	linSVM	76.88	79.89	82.64	85.48	**93.19**	84.08	85.94	88.44	90.11	**95.64**
	polySVM	80.13	82.69	83.09	84.25	90.84	87.58	89.30	89.37	89.80	94.40
	rbfSVM	84.64	85.19	87.00	87.72	87.72	91.41	91.69	92.64	93.01	93.01
	XGBoost	84.31	84.81	86.33	88.26	**92.85**	91.17	90.73	91.31	92.44	**95.34**
	RNN	84.28	85.80	85.50	85.40	91.37	91.12	91.59	90.67	90.45	94.51
	Transformer	84.64	85.19	87.00	85.38	89.72	91.41	91.69	92.64	91.29	94.01

correspond to historical situations in which no comebacks have been observed, and the game outcome is effectively determined by the scoring rules. In these regions, DyWIS primarily acts as a confirmation mechanism that aligns the classifier output with the AW/AL-based decision, rather than uncovering unexpected patterns. Consequently, near-perfect accuracy at late stages reflects structural determinism of the score dynamics combined with the AW/AL thresholds, rather than overfitting or excessive model complexity. As shown in Table 5, Kodai achieves 100% accuracy with DyWIS-XGBoost at $r = 40$, and Viktor attains 100% accuracy with nearly all models from $r = 20$ onwards. Likewise, Kenta, Anders, and Kunlavut exceed 95% with DyWIS-SVM or DyWIS-KNN at $r = 40$. The corresponding F1-scores are also near perfect, which indicates that at this stage, the absolute AW/AL thresholds dominate the decision process, transforming the classification into an almost binary task. Because AW and AL are empirically estimated from historical data, they inevitably reflect not only intrinsic scoring properties but also contextual factors such as opponent distributions, tournament schedules, and periods of dominance or decline in a player's career. In other words, the thresholds are intentionally player-conditioned rather than universal, and they may shift if a player's style, physical condition, or competitive environment changes over time.

These observations are consistent with the transition behavior of DyWIS. Once the ongoing score difference enters the player-specific decisive range, the model transitions from relying on probabilistic ML inference toward confirming the inevitable match outcome. At the same rally stages, both DyWIS-RNN and DyWIS-Transformer converge toward similar near-deterministic predictions as classical models, confirming that once decisive score differences are reached, the influence of the underlying classifier architecture becomes secondary to the threshold-based decision mechanism of DyWIS. However, not all players exhibit identical performance trajectories. An illustrative example is Viktor Axelsen, whose results show several 100% accuracy entries even at $r = 20$ (DyWIS-kNN, DyWIS-RF, DyWIS-rbfSVM, DyWIS-RNN, DyWIS-Transformer). This does not indicate that later rallies contribute more to prediction accuracy but rather reflects the clear dominance patterns in Viktor's matches, which make his outcomes easier to separate. Other classifiers on Viktor (DyWIS-DT, DyWIS-linSVM, DyWIS-XGBoost) do not sustain 100% performance across all rallies. This indicates that the exceptional accuracy is primarily due to the characteristics of his dataset rather than the DyWIS framework itself.

On the other hand, players with more irregular match histories, such as Ng Tze Yong, begin with slightly lower accuracy at early checkpoints of $r = 20$–25, typically ranging around 74–84% depending on the classifier model (Table 5). However, their performance improves markedly at $r = 40$, where nearly all models exceed 90%, and DyWIS-Transformer even attains 96.77%. This indicates that for players whose scoring patterns are more variable, the model needs a longer sequence of rallies to resolve outcome ambiguity. This finding is consistent with the DyWIS mechanism, which means that when the current score difference x_c remains distant from a player's x_w/x_l thresholds, the model relies primarily on the dynamic weighting via w_r and w_s to stabilize intermediate predictions until stronger evidence accumulates.

To further examine the relative contribution of the rule-based AW/AL mechanism and the machine learning classifiers at later rally stages, we analyze the proportion of instances in which the current score difference satisfies the AW or AL threshold. Table 6 summarizes the percentage of condition met at each rally checkpoint, $r = \{20, 25, 30, 35, 40\}$ that fall within these decisive regions for each player. As shown in the table, the proportion of threshold-triggered cases increases consistently as rallies progress. At early stage of $r = 20$, only a small fraction of rallies (typically below 15% for most players) fall within AW/AL regions, which indicating that the machine learning classifiers play a dominant role in prediction. However, as the rally length increases, this proportion rises substantially. At $r = 40$, approximately 25% to 49% of cases across players are already classified directly by the AW/AL thresholds. This trend confirms that the rule-based component contributes increasingly to prediction at later stages, particularly in decisive score regions. Nevertheless, a significant portion of rallies at $r = 40$ still fall outside the AW/AL thresholds, therefore requiring classification by the machine learning models. This demonstrates that DyWIS maintains a hybrid decision structure, where the rule-based logic governs structurally deterministic cases, while machine learning remains essential for resolving ambiguous score states.

TABLE 6. Percentage of rule-based AW/AL conditions met across rally, $r = 20, 25, 30, 35, 40$ by player.

Player	r = 20	r = 25	r = 30	r = 35	r = 40
Kodai	7.89%	13.16%	23.68%	23.68%	23.68%
Jonatan	10.53%	13.16%	28.95%	28.95%	28.95%
ShiYuQi	23.33%	26.67%	30.00%	33.33%	33.33%
LiShiFeng	21.95%	26.83%	36.59%	48.78%	48.78%
Viktor	25.00%	25.00%	25.00%	33.33%	33.33%
Anthony	13.79%	24.14%	31.03%	41.38%	41.38%
Anders	8.33%	14.58%	18.75%	25.00%	25.00%
Kenta	14.49%	20.29%	27.54%	34.78%	34.78%
Kunlavut	29.73%	35.14%	43.24%	45.95%	45.95%
NgTzeYong	9.38%	12.50%	25.00%	25.00%	25.00%

To confirm that these performance gains are not dependent on a specific decision threshold, ROC curves were plotted for all seven DyWIS classifiers at the same rally checkpoints, as shown in Figure 7. At $r = 20$, all models lie well above the diagonal, with AUC values between 0.70 and 0.79, indicating that even early in the game, the system ranks winning rallies above losing ones better than chance. As the number of observed rallies increases, the curves steadily shift toward the ideal point (0, 1), and AUC values rise almost monotonically. At $r = 25$ and $r = 30$, the three best models (DyWIS-KNN, DyWIS-RF and DyWIS-XGBoost)

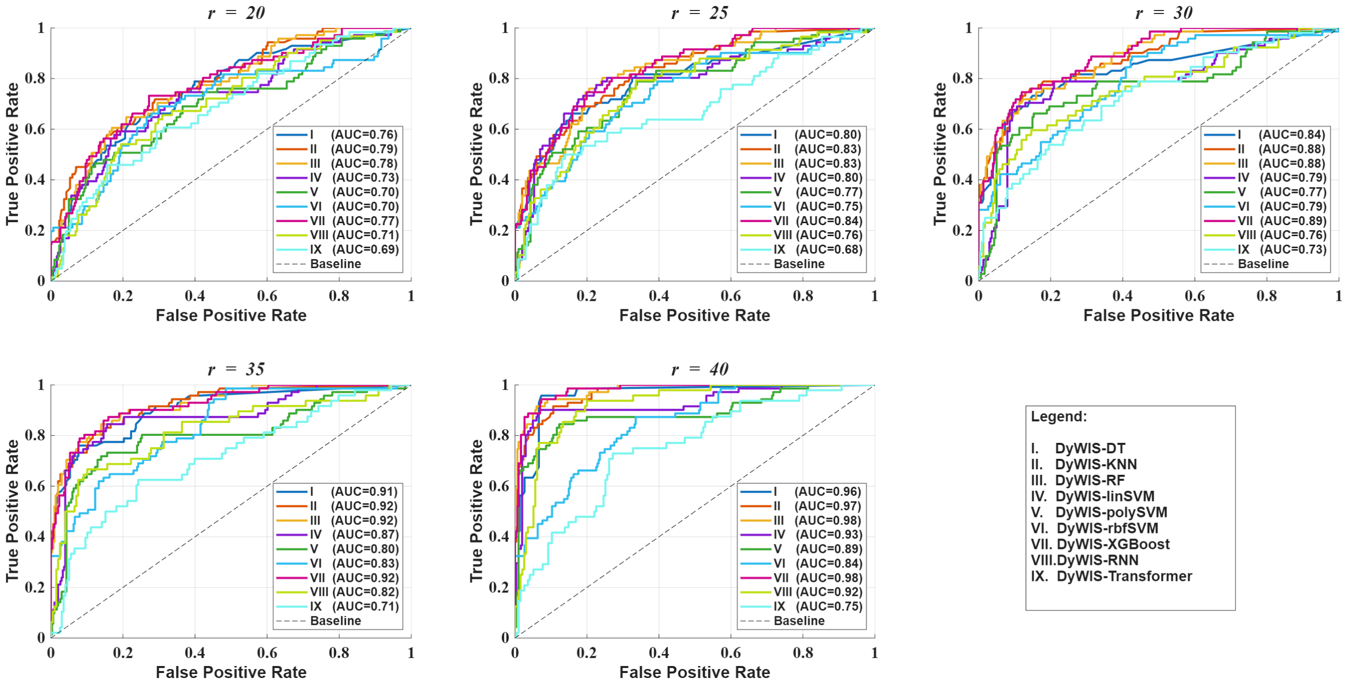


FIGURE 7. ROC curves of the seven DyWIS classifiers at different rally checkpoints $r = \{20, 25, 30, 35, 40\}$.

reach AUCs between 0.83 and 0.89. By $r = 35$ and $r = 40$, most classifier models surpass 0.90, with DyWIS-RF and DyWIS-XGBoost reaching up to 0.98.

This trend aligns closely with the accuracy and F1-score results but provides stronger statistical validation, as ROC analysis evaluates model discrimination across all possible thresholds rather than a single decision point. Across all checkpoints, DyWIS tree-ensemble models and KNN achieve the highest AUCs, followed by linear SVM. The polynomial and RBF SVMs also improve as more rallies are observed but remain slightly below the top-performing group, consistent with the rankings reported in Table 5. Collectively, the ROC curves confirm that later rallies contain richer predictive information, and the proposed DyWIS framework effectively leverages this evolving data without overfitting to the final outcome.

We observed that deep models (RNN and Transformer) do not show a consistent advantage over strong classical approaches in this feature setting. Although their mean accuracies are competitive, they do not surpass the best ensemble or instance-based classifiers at the later rally checkpoints. This indicates that, for live match prediction using score-difference sequences, simpler classifiers within the DyWIS framework can remain highly effective, offering a favourable balance between predictive performance, interpretability, and computational efficiency. Moreover, while KNN is often viewed as memory-intensive in large-scale deployments, its storage cost is modest in our formulation because models are trained per player and each game is represented as a fixed-length sequence. Even after SMOTE balancing, each player's training set contains at most a few dozen game sequences, making the memory footprint of the KNN reference set practical for real-time, match-side analytics.

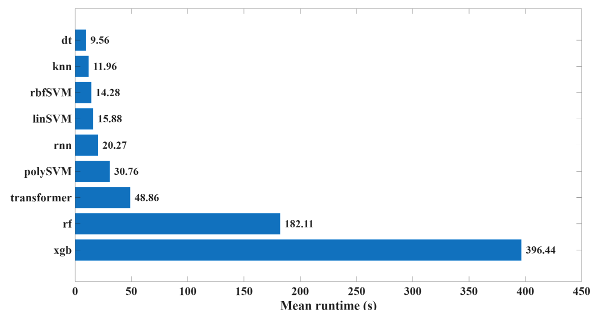


FIGURE 8. Runtime comparison of the evaluated models under the DyWIS framework (mean over 10 datasets).

To further examine computational efficiency, we measured the mean end-to-end runtime at $r = 40$, averaged across the 10 player datasets (Figure 8). The results show that DyWIS-DT, DyWIS-KNN, DyWIS-rbfSVM, and DyWIS-linSVM are substantially faster than DyWIS-RF and DyWIS-XGBoost, while DyWIS-RNN and DyWIS-Transformer fall between these extremes. Finally, the per-player breakdown suggests that predictability is strongly player- and match-context dependent: some players exhibit near-deterministic late-rally outcomes for certain models (e.g., Kodai with XGBoost reaching 100% accuracy at $r = 40$), whereas others show more modest or fluctuating performance across checkpoints. This variability is consistent with differences in comeback likelihood, match closeness, and the possibility of temporal drift in player performance patterns.

Across rally checkpoints, no single model consistently achieves the top accuracy; instead, the best-performing classifier varies with rally length, indicating that the relationship between score-difference features and game outcome evolves during match progression. This suggests that a stage-aware approach by selecting different classifiers for early versus late rallies or an ensemble of complementary models may provide a more robust win-probability estimator than relying on a single classifier. Since the results indicate checkpoint-dependent model dominance, a two-stage strategy can be defined: (i) an early-stage predictor (e.g., RNN) used at r20–r25 where the outcome is less determined, and (ii) a late-stage predictor (e.g., kNN) used at r35–r40 where score-difference states become more decisive. While we do not explicitly evaluate such hybrid selection in this work, the observed checkpoint-dependent ranking motivates future work on rally-stage mixtures-of-experts or stacking ensembles.

To address potential ceiling effects at later rally stages, we additionally performed 3-fold cross-validation on three representative players (Jonatan, Kenta, and Ng Tze Yong) and report accuracy as mean ± standard deviation at $r = 35$ and $r = 40$ (Table 7). Overall, the high mean accuracies at these checkpoints are consistent with increasing determinism as score differences approach player-specific decisive ranges. However, performance is not uniformly perfect across folds, where several models exhibit non-negligible variability, particularly for Ng Tze Yong at $r = 35$, indicating that late-stage accuracy still depends on match variability and the training split. Kenta shows consistently high accuracy with relatively low variance at $r = 40$, whereas Ng Tze Yong exhibits lower and more variable performance at $r = 35$ but improves markedly by $r = 40$, supporting the earlier observation that players with irregular scoring patterns require more rallies for reliable outcome separation. These results suggest that the near ceiling performance at later stages reflects both structural score effects and residual model sensitivity rather than inevitability alone.

B. ABLATION STUDY

To understand how each part of the framework contributes, we conducted an ablation study on the same processed men's singles dataset. Models were trained separately for each player as described in Table 3, with the minimum length of rally, $r = 25$ selected as the standardized input sequence length. The objective was to evaluate how the absolute rule-based layer and the dynamic weighting mechanism individually and jointly affect classification performance. Four model variants were tested: baseline classifier only, absolute win/loss threshold only, dynamic weighting adjustment only, and the full DyWIS model.

All variants were tested on the same dataset and across seven base classifiers (DT, KNN, RF, SVM, XGBoost, RNN and Transformer), as reported in Table 8. The absolute win/loss rule alone achieved an accuracy of only 21.27%. This confirms that the deterministic thresholds by themselves cannot handle intermediate rally states where the score difference is not yet decisive. Similarly, using only the baseline classifier yielded limited predictive accuracy, as the models lacked contextual weighting.

In contrast, incorporating the dynamic weighting adjustment to the ML and DL baseline models led to substantial improvements across all classifiers. For example, accuracy increased from 67.07% to 82.50% for DT, 74.41% to 85.27% for KNN, 67.05% to 80.22% for RF, 67.43% to 84.36% for XGBoost, 67.94% to 85.82% for RNN, and 65.93% to 81.41% for Transformer. These results demonstrate that integrating the historical win rate (w_r) and current score influence (w_s) information is the principal source of performance gains.

The full DyWIS model achieved the highest overall results for most classifiers, with KNN obtaining the best accuracy at 86.59%, followed by RNN at 85.80%. Several models such as RF, RBF SVM, and Transformer achieved same accuracy at 85.19%. A minor exception was observed for DT and RNN, where the ML+dynamic weighting variant (82.50% and 85.82%, respectively) slightly outperformed the full DyWIS (81.96% and 85.80%, respectively). This suggests that the rule-based layer may introduce slight uncertainty, particularly during early rally phases before the score difference becomes conclusive.

C. CASE STUDY: WIN VS. LOSS SCENARIOS

To illustrate how the framework behaves in individual matches rather than only in aggregated statistics, we analyzed two representative games. First, a winning match for Shi Yu Qi game, and second, a losing match from Kodai Naraoka game, as shown in Figure 9. In both plots, we display four rally-wise components, including the predicted win probability from the ML model with dynamic weighting only (without absolute win/loss thresholds), the predicted win probability from the full DyWIS model, the player and opponent scores, and the resulting score difference.

In the winning case (Figure 9(a)), the player gains an early lead, but the score gap does not exceed the absolute win threshold x_w until around the 10th rally. Once the threshold is crossed, DyWIS rapidly pushes the predicted win probability to 1.0, which effectively confirms the victory and then maintains this decision. In contrast, the ML with dynamic weighting increases the probability more gradually and continues updating until the match ends, since there is no rule-based threshold applied. While this behavior may better reflect the continuous evolution of game dynamics, it provides slower decision feedback compared to DyWIS.

The losing case (Figure 9(b)) shows the opposite scenario using the same framework. Here, the player falls behind almost from the start, and the score gap widens over time. Once the score difference reaches the predefined AL threshold, x_l (at the 20th rally), DyWIS immediately drives the predicted win probability toward zero and keeps it there. This rule-based mechanism enables the model to commit early to a confident decision once the match becomes decisively unfavorable. The ML with dynamic weighting

TABLE 7. Cross-validation accuracy (mean ± standard deviation) for DyWIS variants at $r = 35$ and $r = 40$ using 3-fold validation on the Jonatan, Kenta, and Ng Tze Yong datasets.

Classifier	Jonatan		Kenta		NgTzeYong	
	$r = 35$	$r = 40$	$r = 35$	$r = 40$	$r = 35$	$r = 40$
dt	0.947 ± 0.046	0.865 ± 0.126	0.870 ± 0.151	0.913 ± 0.029	0.769 ± 0.157	0.939 ± 0.105
knn	0.974 ± 0.044	0.949 ± 0.089	0.928 ± 0.025	0.971 ± 0.025	0.776 ± 0.108	0.870 ± 0.061
rf	0.974 ± 0.044	0.917 ± 0.144	0.957 ± 0.043	0.971 ± 0.025	0.803 ± 0.105	0.876 ± 0.138
linSVM	0.947 ± 0.046	0.816 ± 0.041	0.913 ± 0.151	0.957 ± 0.043	0.803 ± 0.105	0.906 ± 0.091
polySVM	0.974 ± 0.044	0.923 ± 0.077	0.899 ± 0.025	0.913 ± 0.048	0.775 ± 0.042	0.933 ± 0.058
rbfSVM	0.895 ± 0.043	0.897 ± 0.089	0.899 ± 0.050	0.899 ± 0.050	0.833 ± 0.153	0.836 ± 0.063
XGBoost	0.949 ± 0.044	0.921 ± 0.077	0.913 ± 0.025	0.971 ± 0.025	0.773 ± 0.155	0.906 ± 0.091
rnn	0.893 ± 0.093	0.897 ± 0.118	0.884 ± 0.050	0.899 ± 0.025	0.742 ± 0.052	0.842 ± 0.100
transformer	0.895 ± 0.043	0.921 ± 0.077	0.899 ± 0.025	0.901 ± 0.050	0.869 ± 0.061	0.842 ± 0.141

TABLE 8. Ablation study on the performance of proposed models and base machine learning models.

Method	Mean accuracy (%)									
	DT	KNN	RF	linSVM	polySVM	rbfSVM	XGB	RNN	Transformer	Rule-based
Baseline ML Classifier	67.07	74.41	67.05	68.12	62.92	33.98	67.43	67.94	65.93	—
Absolute Win/Loss Threshold (Proposed Model)	—	—	—	—	—	—	—	—	—	21.27
ML/DL + Dynamic Weighting Adjustment (Proposed Model)	82.50	85.27	80.22	78.60	80.54	80.22	84.36	85.82	81.41	—
DyWIS (Proposed Model)	81.96	86.59	85.19	79.89	82.69	85.19	84.81	85.80	85.19	—

adjustment model follows a similar downward trend but continues to update even after the outcome is effectively clear.

In practice, this early decision-making capability of DyWIS is valuable for real-time applications such as live match analytics or in-play betting, where fast and reliable updates are critical. This is also the main reason we retain the Absolute Win/Loss Threshold component within the DyWIS framework, even though using it in isolation yields relatively low overall accuracy (see Table 8). It provides the crucial mechanism that allows DyWIS to issue timely, confident predictions once the game trend becomes decisive.

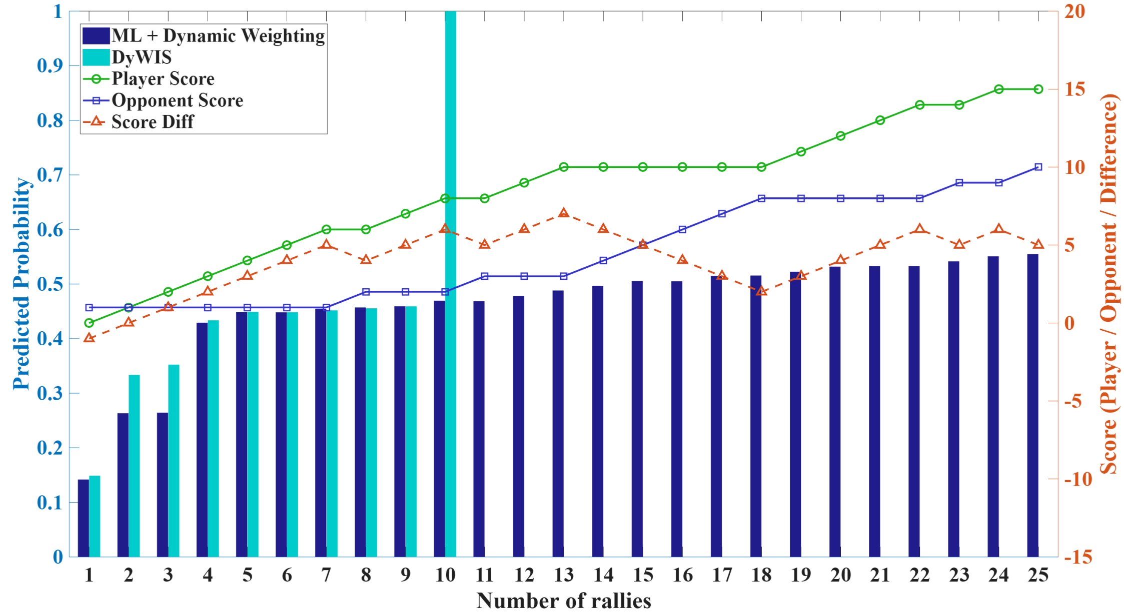
D. COMPARISON WITH PREVIOUS WORKS

After analysing the DyWIS variants and their rally-by-rally behaviour, we next compare the proposed framework with existing badminton and other general sports outcome models. Table 9 reports the results of three representative studies that we reimplemented on the same dataset (Yuan et al. [35], Hodge et al. [47], and Roumani et al. [63]) alongside the DyWIS variants. These results are evaluated at the rally checkpoint of $r = 40$.

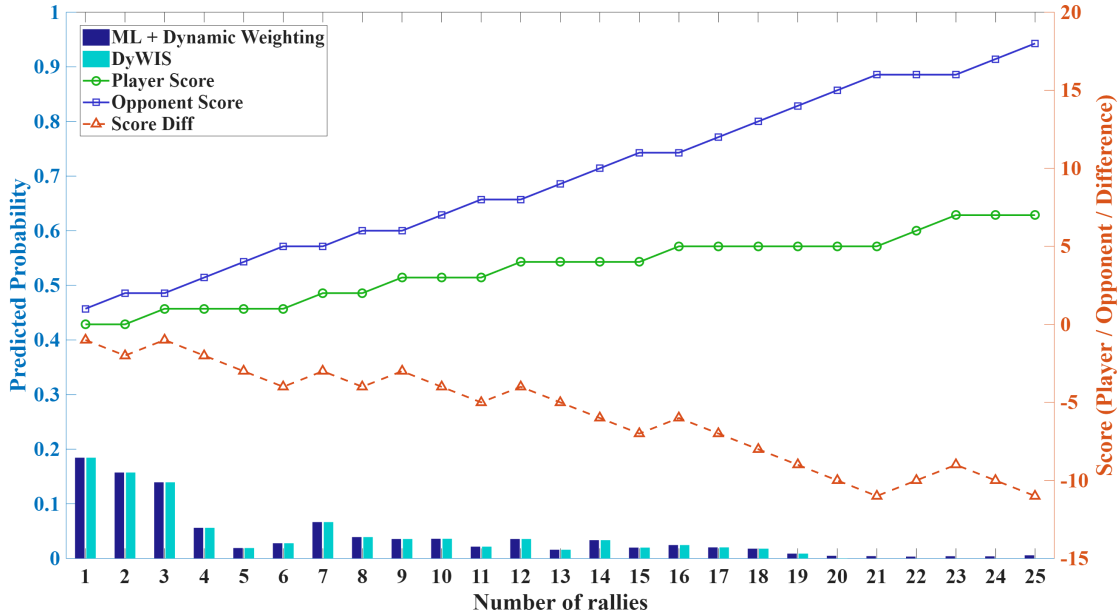
Table 9 reports the accuracies for ten players as well as the mean value across players. Since all models were evaluated on the same preprocessed matches, the observed performance differences can be attributed to the modelling strategy rather than to data variation. The methods from Yuan et al. [35] form the lower band of performance, with mean accuracies ranging from 80.97% (DT) to 85.98% (XGBoost). Both Hodge et al. [47] and Roumani et al. [63] achieve higher accuracies than Yuan et al. In particular, Hodge et al. report 87.74% with RF and 89.12% with Gradient Boosting, while Roumani et al. obtain 86.16% with RF. These results confirm that tree-based ensemble models are effective for rally-outcome prediction, even when they were not originally designed specifically for badminton.

The proposed DyWIS framework exceeds this performance range for most classifiers. DyWIS-KNN achieves the highest mean accuracy at 93.37%, followed closely by DyWIS-linSVM at 93.19% and DyWIS-XGBoost at 92.85%. For sequence-based variants, DyWIS-RNN achieves 91.37% mean accuracy, which is competitive but still below the strongest conventional DyWIS classifiers, while DyWIS-Transformer achieves 89.72%, remaining above most prior baselines but not matching the top DyWIS configurations. The DyWIS-polySVM variant also performs strongly with 90.84%, outperforming all methods from Yuan et al. and Roumani et al. Even the DyWIS-DT reaches 89.39%, clearly above the decision-tree baselines (80.97% in [35] and 84.69% in [63]). The only case where the improvement is marginal is within the random forest family. The DyWIS-RF and DyWIS-rbfSVM both achieve 87.72%, which is effectively comparable to the RF reported by Hodge et al. (87.74%). This suggests that the RF-based models may already be operating near close to their performance limit on this dataset.

A similar pattern is observed at the player level. For athletes with more irregular match histories, such as Jonatan Christie, Shi Yu Qi, and Ng Tze Yong, baseline methods in Table 9 often remain below 88%, whereas DyWIS variants routinely exceed 90% and in several cases surpass 93% (e.g. DyWIS-KNN and DyWIS-linSVM). Notably, Ng Tze Yong reaches 96.77% accuracy with DyWIS-Transformer, which indicates that sequence-based models can be especially effective for specific player even though they do not dominate the mean ranking across all players. This indicates that the proposed framework is more effective at exploiting rally-by-rally contextual information, in contrast to the earlier studies that typically generate a single prediction from a static feature vector. By reweighting the classifier output using the



(a)



(b)

FIGURE 9. Progression of predicted winning probabilities from rally 1 to rally 25 for DyWIS and the ML model with dynamic weighting together with score evolution. (a) Winning case: Shi Yu Qi's game (b) Losing case: Kodai Naraoka's game.

current score difference and player-specific absolute win/loss thresholds, DyWIS stabilizes the prediction once the match enters a decisive phase.

Overall, Table 9 shows that previous methods provide competitive baselines in the 81–89% range, yet the proposed DyWIS framework consistently lifts the best accuracies into the 92–93% range across multiple classifiers. This confirms that updating the model with the current rally score and integrating absolute thresholds makes live badminton outcome prediction more accurate and robust than conventional static-feature approaches.

While DyWIS operates purely on rally-level score differences, this representation primarily captures macro-level properties of play such as momentum shifts, dominance patterns, and resilience under pressure when trailing or protecting a lead. It does not directly capture micro-level

TABLE 9. Comparison performances of mean accuracy between proposed models and existing works.

Study	Classifier	Kodai	Jonatan	ShiYuQi	LiShiFeng	Viktor	Anthony	Anders	Kenta	Kunlavut	NgTzeYong	Mean
Yuan et al. [35]	DT	82.76	67.86	77.27	74.19	77.78	85.71	97.22	86.27	81.48	79.17	80.97
	KNN	82.76	82.14	77.27	70.97	88.89	90.48	97.22	90.20	77.78	75.00	83.27
	RF	93.10	78.57	68.18	77.42	100.00	90.48	97.22	88.24	85.19	79.17	85.76
	rbfSVM	82.76	85.71	72.73	74.19	77.78	90.48	97.22	92.16	81.48	79.17	83.37
	XGBoost	89.66	85.71	68.18	80.65	88.89	90.48	97.22	86.27	85.19	87.50	85.98
Hodge et al. [47]	RF	93.18	86.36	88.24	85.11	71.43	90.91	96.36	92.41	92.86	80.56	87.74
	GB	93.18	84.09	88.24	87.23	78.57	87.88	94.55	94.94	88.10	94.44	89.12
Roumani et al. [63]	RF	93.10	78.57	72.73	80.65	100.00	90.48	97.22	88.24	81.48	79.17	86.16
	DT	86.21	75.00	77.27	77.42	88.89	90.48	94.44	88.24	81.48	87.50	84.69
DyWIS (this work)	DT	92.11	86.84	93.33	85.37	83.33	85.71	95.83	91.30	86.49	93.55	89.39
	KNN	97.37	92.11	86.67	87.80	100.00	92.86	95.83	95.65	91.89	93.55	93.37
	RF	84.21	89.47	90.00	87.80	100.00	82.14	83.33	89.86	86.49	83.87	87.72
	linSVM	97.37	84.21	86.67	92.68	100.00	89.29	97.92	95.65	94.59	93.55	93.19
	polySVM	86.84	92.11	86.67	82.93	100.00	89.29	91.67	91.30	97.30	90.32	90.84
	rbfSVM	84.21	89.47	90.00	87.80	100.00	82.14	83.33	89.86	86.49	83.87	87.72
	XGBoost	100.00	84.21	93.33	85.37	100.00	89.29	93.75	97.10	91.89	93.55	92.85
	RNN	92.11	92.11	93.33	90.24	100.00	89.29	89.58	91.30	91.89	83.87	91.37
	Transformer	84.21	89.47	90.00	87.80	100.00	89.29	83.33	89.86	86.49	96.77	89.72

tactical decisions e.g., stroke selection, shot combinations, or spatial formations, nor physiological states such as fatigue or recovery. As such, our framework should be viewed as complementary to spatiotemporal or video-based models that explicitly encode technical and positional information.

V. CONCLUSION

This work presented a Dynamic Weighted Influence Score (DyWIS) framework for in-game win probability prediction in elite badminton men's singles. Unlike conventional approaches that rely on static features or computationally expensive video-based inputs, the proposed method exploits rally-level score differences, combined with player-specific absolute win/loss (AW/AL) thresholds and dynamic weighting based on historical win rates and current score progression. Using a large-scale BWF dataset of top-10 players, DyWIS was integrated with several machine learning classifiers and evaluated at multiple rally checkpoints to reflect realistic live-match scenarios.

Experimental results demonstrated that DyWIS consistently improves performance over baseline classifiers and prior badminton outcome models. At early stages (around $r = 20$ rallies), the framework already delivers highly informative predictions, with many player–classifier combinations exceeding 80% accuracy and strong F1-scores, enabling meaningful tactical decisions before a game reaches its minimum deciding rally. As rallies accumulate, prediction accuracy, F1-score, and AUC continue to increase, and for several players the performance approaches near-deterministic levels by $r = 40$. The ablation study showed that the dynamic weighting component through historical win rate and score influence is the main driver of performance gains, while the AW/AL rule-based layer provides a mechanism for early, confident decisions once the score enters decisive regions. In comparative experiments, DyWIS-KNN, DyWIS-linSVM, and DyWIS-XGBoost variants achieved mean accuracies of 93.37%, 93.19%, and 92.85%, respectively, outperforming reimplemented baselines from the literature across most classifiers and players.

Despite these encouraging results, several limitations remain. The current study focuses on top-10 men's singles players, which constrains generalizability to broader ranking levels, other categories (e.g., women's singles, doubles), and junior or amateur competitions. Extension to women's singles would require recalibration of player-specific AW/AL thresholds and historical win-rate distributions. For doubles categories, the framework could be adapted by redefining score-difference trajectories at the pair level and incorporating pair-specific historical performance. Moreover, the player-specific AW/AL thresholds are derived from historical matches pooled across the 2018–2023 study period, which may not fully capture temporal variations in player form, dominance, or comeback ability across seasons. Time-aware threshold estimation, such as season-specific or rolling-window AW/AL updates could improve sensitivity of DyWIS to evolving player performance. Furthermore, although SMOTE was used to handle class imbalance, it does not explicitly preserve the sequential nature of rally data progression. This is because SMOTE interpolates between samples in feature space without enforcing the sequential constraints of actual badminton score development. While this is partly consistent with the current DyWIS formulation, sequence-aware augmentation strategies may provide more realistic synthetic training samples in future work.

Nonlinear models also showed sensitivity to irregular playing styles, suggesting that more advanced sequence-aware architectures could further enhance robustness. Future work will explore sequence-preserving and multimodal extensions, such as incorporating stroke type, spatial information, physiological indicators, and momentum-sensitive features based on the rate of change of score difference over recent rallies to better capture short-term momentum shifts. In addition, contextual match variables such as serving side may provide further tactical and psychological information, which could improve the sensitivity of DyWIS to in-game state changes while preserving its real-time applicability. Also, the current DyWIS framework is strictly match-internal and does not explicitly account for external load factors such as cumulative fatigue from tournament scheduling. Since prior badminton research has shown that participation load can affect performance consistency, incorporating schedule-related fatigue indicators may further improve rally weighting in future work.

From a practical perspective, the proposed DyWIS framework is lightweight, interpretable, and suitable for real-time deployment. Its ability to update win probabilities after each rally and to lock in outcomes once decisive thresholds are reached makes it particularly attractive for coaching dashboards, performance analysis tools, broadcast augmentation, and in-play decision support. Overall, this study demonstrates that integrating score-difference dynamics with dynamic weighting and absolute thresholds offers a powerful and generalizable approach to in-game win probability prediction in elite badminton.

ACKNOWLEDGMENT

The authors would like to thank to Universiti Putra Malaysia (UPM) for the support in terms of resources and facilities provided during this research.

REFERENCES

[1] K. Kapadia, H. Abdel-Jaber, F. Thabtah, and W. Hadi, "Sport analytics for cricket game results using machine learning: An experimental study," *Appl. Comput. Informat.*, vol. 18, no. 3/4, pp. 256–266, Jun. 2022, doi: 10.1016/j.aci.2019.11.006.

[2] Y. Nokihara, R. Hachiuma, R. Hori, and H. Saito, "Future prediction of shuttlecock trajectory in badminton using player's information," *J. Imag.*, vol. 9, no. 5, p. 99, May 2023, doi: 10.3390/jimaging9050099.

[3] M. Oytun, C. Tinazci, B. Sekeroglu, C. Acikada, and H. U. Yavuz, "Performance prediction and evaluation in female handball players using machine learning models," *IEEE Access*, vol. 8, pp. 116321–116335, 2020, doi: 10.1109/ACCESS.2020.3004182.

[4] F. R. Goes, L. A. Meerhoff, M. J. O. Bueno, D. M. Rodrigues, F. A. Moura, M. S. Brink, M. T. Elferink-Gemser, A. J. Knobbe, S. A. Cunha, R. S. Torres, and K. A. P. M. Lemmink, "Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review," *Eur. J. Sport Sci.*, vol. 21, no. 4, pp. 481–496, Apr. 2021, doi: 10.1080/17461391.2020.1747552.

[5] C. Xu, "Tactical intelligent decision modelling in sports competitions based on reinforcement learning algorithms," *J. Electr. Syst.*, vol. 20, no. 6s, pp. 2092–2101, Apr. 2024, doi: 10.52783/jes.3124.

[6] L. Zhang and Y. An, "Enhancing sports team management through machine learning," *IEEE Access*, vol. 13, pp. 55431–55441, 2025, doi: 10.1109/ACCESS.2025.3551889.

[7] A. A. Phatak, S. Mehta, F.-G. Wieland, M. Jamil, M. Connor, M. Bassek, and D. Memmert, "Context is key: Normalization as a novel approach to sport specific preprocessing of KPI's for match analysis in soccer," *Sci. Rep.*, vol. 12, no. 1, pp. 1–6, Jan. 2022, doi: 10.1038/s41598-022-05089-y.

[8] Z. Liu, X. Xie, M. He, W. Zhao, Y. Wu, L. Cheng, H. Zhang, and Y. Wu, "Smartboard: Visual exploration of team tactics with LLM agent," *IEEE Trans. Vis. Comput. Graphics*, vol. 31, no. 1, pp. 23–33, Jan. 2025, doi: 10.1109/TVCG.2024.3456200.

[9] H. E. Sinadia and I. M. Murwantara, "Sports analytics: A comparison of machine learning perfomance for profiling badminton athlete," in *Proc. 1st Int. Conf. Technol. Innov. Its Appl. (ICTIIA)*, Tangerang, Indonesia, Sep. 2022, pp. 1–6, doi: 10.1109/ICTIIA54654.2022.9935852.

[10] H. McGuigan, P. Hassmén, N. Rosic, and C. J. Stevens, "Training monitoring methods used in the field by coaches and practitioners: A systematic review," *Int. J. Sports Sci. Coaching*, vol. 15, no. 3, pp. 439–451, Jun. 2020, doi: 10.1177/1747954120913172.

[11] R. Bunker and T. Susnjak, "The application of machine learning techniques for predicting match results in team sport: A review," *J. Artif. Intell. Res.*, vol. 73, pp. 1285–1322, Apr. 2022, doi: 10.1613/jair.1.13509.

[12] H. Van Eetvelde, L. D. Mendonça, C. Ley, R. Seil, and T. Tischer, "Machine learning methods in sport injury prediction and prevention: A systematic review," *J. Experim. Orthopaedics*, vol. 8, no. 1, pp. 1–15, Jan. 2021, doi: 10.1186/s40634-021-00346-x.

[13] A. Wong, E. Li, H. Le, G. Bhangu, and S. Bhatia, "A predictive analytics framework for forecasting soccer match outcomes using machine learning models," *Decis. Analytics J.*, vol. 14, Mar. 2025, Art. no. 100537, doi: 10.1016/j.dajour.2024.100537.

[14] Z. Chen, "A hybrid machine learning framework for soccer match outcome prediction: Incorporating bivariate Poisson distribution," *Proc. ITM Web Conf.*, vol. 70, p. 03020, Jan. 2025.

[15] H. Arntzen and L. M. Hvattum, "Predicting match outcomes in association football using team ratings and player ratings," *Stat. Model.*, vol. 21, no. 5, pp. 449–470, Oct. 2021, doi: 10.1177/1471082x20929881.

[16] Z. Gao and A. Kowalczyk, "Random forest model identifies serve strength as a key predictor of tennis match outcome," *J. Sports Analytics*, vol. 7, no. 4, pp. 255–262, Dec. 2021, doi: 10.3233/jsa-200515.

[17] M.-L. Huang and Y.-Z. Li, "Use of machine learning and deep learning to predict the outcomes of major league baseball matches," *Appl. Sci.*, vol. 11, no. 10, p. 4499, May 2021, doi: 10.3390/app11104499.

[18] I. Kekez, M. Ćukušić, and M. Jadrić, "Data mining approach for business value analysis in basketball," *Zbornik Veleučilišta u Rijeci*, vol. 9, no. 1, pp. 227–248, 2021, doi: 10.31784/zvr.9.1.14.

[19] V. Sarlis and C. Tjortjis, "Sports analytics: Data mining to uncover NBA player position, age, and injury impact on performance and economics," *Inf. (Switzerland)*, vol. 15, no. 4, p. 242, 2024, doi: 10.3390/info15040242.

[20] T. Sasikala, J. J. D. Raj, B. Swathi, S. Afreen, and K. Naveen, "World cup cricket data analytics: A comparative review of T20 and ODI world cup prediction techniques," in *Proc. Int. Conf. Knowl. Eng. Commun. Syst. (ICKECS)*, Chickballapur, India, Apr. 2025, pp. 1–6, doi: 10.1109/ickecs65700.2025.11034937.

[21] R. Elimam, N. Sutton-Charani, J. Prioux, J. Montmain, and S. Perrey, "Multi-output regression for the prediction of world-class performances in women's handball," *IEEE Access*, vol. 13, pp. 69873–69887, 2025, doi: 10.1109/ACCESS.2025.3560838.

[22] J. Wu, D. Liu, Z. Guo, Q. Xu, and Y. Wu, "TacticFlow: Visual analytics of ever-changing tactics in racket sports," *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 1, pp. 835–845, Jan. 2022, doi: 10.1109/TVCG.2021.3114832.

[23] P.-E. Martin, J. Benois-Pineau, and R. Péteri, "Fine-grained action detection and classification in table tennis with Siamese spatio-temporal convolutional neural network," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Taipei, Taiwan, Sep. 2019, pp. 3027–3028, doi: 10.1109/ICIP.2019.8803382.

[24] M. A. King, H. Towler, R. N. Dillon, and S. A. McErlain-Naylor, "A correlational analysis of shuttlecock speed kinematic determinants in the badminton jump smash," *Appl. Sci. (Switzerland)*, vol. 10, no. 4, p. 1248, 2020, doi: 10.3390/app10041248.

[25] V. T. R. P. Kumar, S. Imran, S. K. Masthan, O. P. Kumar, S. R. Sulthana, and V. Yuktasha, "Data-driven decision making in sports: Machine learning for performance benchmarking," in *Proc. 3rd Int. Conf. Augmented Intell. Sustain. Syst. (ICAISS)*, Trichy, India, May 2025, pp. 1469–1476, doi: 10.1109/icaiss61471.2025.11042003.

[26] N. Ding, K. Takeda, and K. Fujii, "Deep reinforcement learning in a racket sport for player evaluation with technical and tactical contexts," *IEEE Access*, vol. 10, pp. 54764–54772, 2022, doi: 10.1109/ACCESS.2022.3175314.

[27] Z. Guo, L. Jiang, C. Liu, Z. Yang, S. Qin, D. Jiang, D. Bao, J. Dai, and H. Liu, "Using inertial measurement units to quantify external load in men's singles badminton matches: Insights from set outcomes and score gaps," *J. Sports Sci. Med.*, vol. 24, no. 4, pp. 901–909, Nov. 2025, doi: 10.52082/jssm.2025.901.

[28] K. C. Houtmeyers, A. Jaspers, and P. Figueiredo, "Managing the training process in elite sports: From descriptive to prescriptive data analytics," *Int. J. Sports Physiol. Perform.*, vol. 16, no. 11, pp. 1719–1723, Nov. 2021, doi: 10.1123/ijspp.2020-0958.

[29] M. Ibh, S. Graßhof, and D. W. Hansen, "A stroke of genius: Predicting the next move in badminton," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2024, pp. 3376–3385.

[30] W. Wang, H.-H. Shuai, K.-S. Chang, and W.-C. Peng, "ShuttleNet: Position-aware fusion of rally progress and player styles for stroke forecasting in badminton," in *Proc. 36th AAAI Conf. Artif. Intell.*, 2021, pp. 4219–4227, doi: 10.48550/arxiv.2112.01044.

[31] M. Á. Gómez-Ruano, A. Cid, F. Rivas, and L. M. Ruiz, "Serving patterns of Women's badminton medalists in the rio 2016 Olympic games," *Frontiers Psychol.*, vol. 11, pp. 1–9, Feb. 2020, doi: 10.3389/fpsyg.2020.00136.

[32] P. Liu and J.-H. Wang, "MonoTrack: Shuttle trajectory reconstruction from monocular badminton video," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, New Orleans, LA, USA, Jun. 2022, pp. 3512–3521, doi: 10.1109/CVPRW56347.2022.00395.

[33] M. Sharma, Monika, N. Kumar, and P. Kumar, "Badminton match outcome prediction model using Naıve Bayes and feature weighting technique," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 8, pp. 8441–8455, Aug. 2021, doi: 10.1007/s12652-020-02578-8.

[34] Y. Sheng, C. Liu, Q. Yi, W. Ouyang, R. Wang, and P. Chen, "Predicting badminton outcomes through machine learning and technical action frequencies," *Sci. Rep.*, vol. 15, no. 1, pp. 1–8, Mar. 2025, doi: 10.1038/s41598-025-87610-7.

[35] H. Yuan, Y. Wang, K. Yang, and Y. Bin, "Prediction model and technical and tactical decision analysis of women's badminton singles based on machine learning," *PLoS ONE*, vol. 19, no. 11, pp. 1–20, Nov. 2024, doi: 10.1371/journal.pone.0312801.

[36] Y. Jia, N. A. Abdullah, H. Eliza, Q. Lu, D. Si, H. Guo, and W. Wang, "A narrative review of deep learning applications in sports performance analysis: Current practices, challenges, and future directions," *BMC Sports Sci.*, vol. 17, no. 1, p. 249, 2025, doi: 10.1186/s13102-025-01294-0.

[37] J. Mohana, R. Bharthvajan, M. Murugesan, H. Omotunde, D. Balaji, and A. S. M. Antony, "An AI-driven approach to real-time sports analytics and performance evaluation," in *Proc. Int. Conf. Innov. Comput., Intell. Commun. Smart Electr. Syst. (ICSES)*, Chennai, India, Dec. 2024, pp. 1–7, doi: 10.1109/icses63760.2024.10910720.

[38] T. Lin, A. Aouididi, C. Zhu-Tian, J. Beyer, H. Pfister, and J.-H. Wang, "VIRD: Immersive match video analysis for high-performance badminton coaching," *IEEE Trans. Vis. Comput. Graphics*, vol. 30, no. 1, pp. 458–468, Jan. 2024, doi: 10.1109/TVCG.2023.3327161.

[39] T. Zhao, J. Cabral, and G. Zhu, "A novel explainable artificial intelligence framework using knockoffs techniques with applications to sports analytics," *Ann. Operations Res.*, 2025, doi: 10.1007/s10479-025-06575-y.

[40] B. Wilczyński, M. Biały, and K. Zorena, "An interpretable machine learning framework for athlete motor profiling using multi-domain field assessments: A proof-of-concept study," *Appl. Sci.*, vol. 15, no. 12, p. 6436, Jun. 2025, doi: 10.3390/app15126436.

[41] Q. Li, T.-C. Chiu, H.-W. Huang, M.-T. Sun, and W.-S. Ku, "VideoBadminton: A video dataset for badminton action recognition," in *Proc. IEEE Int. Conf. Big Data (BigData)*, Washington, DC, USA, Dec. 2024, pp. 1387–1392, doi: 10.1109/BIGDATA62323.2024.10825009.

[42] D. Krstić, T. Vučković, D. Dakić, S. Ristić, and D. Stefanović, "The application and impact of artificial intelligence on sports performance improvement: A systematic literature review," in *Proc. 4th Int. Conf. Commun., Inf., Electron. Energy Syst. (CIEES)*, Plovdiv, Bulgaria, Nov. 2023, pp. 1–8, doi: 10.1109/ciees58940.2023.10378750.

[43] F. Felice and C. Ley, "Predicting handball matches with machine learning and statistically estimated team strengths," *J. Sports Analytics*, vol. 11, pp. 1–12, Jan. 2025, doi: 10.1177/22150218251313937.

[44] J.-M. Hung, J.-Y. Chiang, and K. Wang, "Tennis player pose classification using YOLO and MLP neural networks," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Hualien City, Taiwan, Nov. 2021, pp. 1–2, doi: 10.1109/ISPACS51563.2021.9650925.

[45] S. Çıklaçandır, İ. Kanat, F. G. Y. Çıklaçandır, and Ö. K. Cura, "Classification of tennis serve according to players' experience through wearable motion capture technology," in *Proc. Med. Technol. Congr. (TIPTEKNO)*, 2024, pp. 1–4, doi: 10.1109/tiptekno63488.2024.10755243.

[46] R. Metulini and G. Gnecco, "Measuring players' importance in basketball using the generalized Shapley value," *Ann. Operations Res.*, vol. 325, no. 1, pp. 441–465, Jun. 2023, doi: 10.1007/s10479-022-04653-z.

[47] V. J. Hodge, S. Devlin, N. Sephton, F. Block, P. I. Cowling, and A. Drachen, "Win prediction in multiplayer esports: Live professional match prediction," *IEEE Trans. Games*, vol. 13, no. 4, pp. 368–379, Dec. 2021, doi: 10.1109/TG.2019.2948469.

[48] C. Karunakaran, V. Niranjan, and A. S. Setlur, "Random forest and XGBoost-based ensemble models for colorectal cancer exome variant classification and web application deployment for early prediction," *Comput. Structural Biotechnol. J.*, vol. 21, Apr. 2023, Art. no. 100063, doi: 10.1016/j.csbr.2025.100063.

[49] D. Huang, Z. Liu, and D. Wu, "Research on ensemble learning-based feature selection method for time-series prediction," *Appl. Sci.*, vol. 14, no. 1, p. 40, 2023, doi: 10.3390/app14010040.

[50] C. S. Ranganathan, P. Pandey, M. Arulprakash, K. Gopalakrishnan, T. R. GaneshBabu, and S. Murugan, "Artificial neural networks for enhancing soccer team performance through tactical data analysis," in *Proc. Int. Conf. Mach. Learn. Auto. Syst. (ICMLAS)*, Prawet, Thailand, Mar. 2025, pp. 642–647, doi: 10.1109/icmlas64557.2025.10969039.

[51] Y.-H. Hsu, C.-C. Yu, and H.-Y. Cheng, "Enhancing badminton game analysis: An approach to shot refinement via a fusion of shuttlecock tracking and hit detection from monocular camera," *Sensors*, vol. 24, no. 13, p. 4372, 2024, doi: 10.3390/s24134372.

[52] M.-Á. Gómez, F. Rivas, A. S. Leicht, and J. M. Buldú, "Using network science to unveil badminton performance patterns," *Chaos, Solitons Fractals*, vol. 135, Jun. 2020, Art. no. 109834, doi: 10.1016/j.chaos.2020.109834.

[53] D. Hoffmann, M. Brinkschulte, and T. Vogt, "Match characteristics of women's and men's singles and doubles events in elite badminton," *Int. J. Perform. Anal. Sport*, vol. 26, no. 1, pp. 1–18, Jan. 2026, doi: 10.1080/24748668.2024.2386830.

[54] J.-H. Hsu, H.-L. Lin, H.-C. Fan Chiang, D.-S. Lee, Y. Lee, C.-W. Huang, and Z.-F. Yao, "Optimizing performance in badminton tournaments: The relationship between timing, quantity, and quality among professional players," *J. Funct. Morphology Kinesiol.*, vol. 10, no. 1, p. 5, Dec. 2024, doi: 10.3390/jfmk10010005.

[55] W. Gawin, C. Beyer, and M. Seidler, "A competition analysis of the single and double disciplines in world-class badminton," *Int. J. Perform. Anal. Sport*, vol. 15, no. 3, pp. 997–1006, Dec. 2015, doi: 10.1080/24748668.2015.11868846.

[56] W. H. M. Isa, M. A. Abdullah, M. A. M. Razman, A. P. P. A. Majeed, and I. M. Khairuddin, "Deep learning algorithms for recognition of badminton strokes: A study using SDNN, RNN, and RNN-GRU models with off-court video capture," in *Advances in Intelligent Manufacturing and Robotics*. Cham, Switzerland: Springer, 2024, pp. 53–60.

[57] C. Zong, "Research on badminton movement machine learning model based on computer vision technology," in *Proc. IEEE 2nd Int. Conf. Image Process. Comput. Appl. (ICIPCA)*, vol. 6, Jun. 2024, pp. 550–554, doi: 10.1109/ICIPCA61593.2024.10709287.

[58] I. Farady, Z.-H. Lee, and C.-Y. Lin, "A real-time deep learning approach to badminton player positioning and movement tracking," in *Proc. IEEE Int. Conf. Adv. Vis. Signal-Based Syst. (AVSS)*, Tainan, Taiwan, Aug. 2025, pp. 1–6, doi: 10.1109/AVSS65446.2025.11149910.

[59] Y. E. Tan, J. See, J. Abdullah, L. K. Wang, and K. W. Ban, "A deep learning based framework for badminton rally outcome prediction," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Penang, Malaysia, Nov. 2022, pp. 1–5, doi: 10.1109/ISPACS57703.2022.10082764.

[60] M. A. Haq, S. Tarashima, and N. Tagawa, "Bag of tricks toward enhanced shuttlecock detection from badminton videos," in *Proc. 12th Int. Electr. Eng. Congr. (iEECON)*, Mar. 2024, pp. 1–6, doi: 10.1109/iEECON60677.2024.10537809.

[61] C. Richter, E. King, S. Strike, and A. Franklyn-Miller, "Objective classification and scoring of movement deficiencies in patients with anterior cruciate ligament reconstruction," *PLoS ONE*, vol. 14, no. 7, Jul. 2019, Art. no. e0206024, doi: 10.1371/journal.pone.0206024.

[62] Badminton World Federation. *BWF World Tour*. Accessed: Nov. 15, 2025. [Online]. Available: https://bwfworldtour.bwfbadminton.com

[63] Y. F. Roumani, "Sports analytics in the NFL: Classifying the winner of the superbowl," *Ann. Operations Res.*, vol. 325, no. 1, pp. 715–730, Jun. 2023, doi: 10.1007/s10479-022-05063-x.

NUR ATHIRAH MOHD SHAZALI received the bachelor's degree (Hons.) in instrumentation science from Universiti Putra Malaysia (UPM), in 2022, where she is currently pursuing the master's degree with the Faculty of Science. She has experience in web scraping, data pre-processing, pattern recognition techniques, and building supervised learning models using tools, such as Python and MATLAB. Her research interests include data analytics, machine learning algorithm, sports analytics, and the application of artificial intelligence in real-world systems.

MOHD AMIRUDDIN ABD RAHMAN (Member, IEEE) received the B.Sc. degree in electrical engineering from Purdue University and the M.Sc. degree in sensor and instrumentation from Universiti Putra Malaysia (UPM). He is currently an Associate Professor of data science with the Department of Physics, UPM. He is also the Deputy Director of the University's Research Management Centre. He is a Pioneer in the co-tutelle Ph.D. Program, having earned the first such degree from both the University of Sheffield and UPM. His research experience extends globally, including stints as an Invited Researcher with Alcatel-Lucent Bell Laboratories and a Postdoctoral Fellow with Indian Institute of Technology Kanpur. He has also played a pivotal role in national research initiatives, contributing significantly to the Smart Data and Big Data Projects at the Ministry of Higher Education. He leads the Computational Intelligence Research Group at UPM, and his research focuses on machine learning and artificial intelligence, with more than 50 high-impact publications in the field. His research interests include signal processing, pattern recognition, machine learning algorithms, electromagnetics-related computation and modeling, and RF and microwave-based sensor systems.

MUHAMMAD LUQMAN ARIF MOHAMAD received the bachelor's degree (Hons.) in physics, in 2022. He is currently pursuing the Ph.D. degree with the Faculty of Science, Universiti Putra Malaysia. Additionally, he participated in the notable Taiwan's National Science and Technology Council (NSTC) International Internship Pilot Program (IIPP). His research interests include deep learning algorithms for computer vision, particularly exploring transformer and convolutional neural networks for video action recognition. He was a recipient of the prestigious MyBrainSc government scholarship for his doctoral studies.

ATHIYYAH QISTINA ABDUL HAKIM received the bachelor's degree (Hons.) in physics from Universiti Putra Malaysia (UPM), Serdang, in 2023, where she is currently pursuing the master's degree with the Department of Physics. She has worked with various predictive modeling techniques, such as ARIMA, long short-term memory (LSTM), recurrent neural networks (RNN), and support vector regression (SVR) in the context of supply chain and market price forecasting. Her research interests include time series forecasting, machine learning, and agricultural commodity price prediction.

CACEJA ELYCA ANAK BUNDAK received the B.S. degree (Hons.) in major physics and the master's degree in applied physics from Universiti Putra Malaysia (UPM), in 2018. She is currently pursuing the Ph.D. degree in data science. Her research interests include data analytics, machine learning, and indoor localization systems.

SITI FAIRUZ MAT RADZI received the Bachelor of Science degree (Hons.) in applied physics from Universiti Tun Hussein Onn Malaysia, in 2019, and the Master's of Science degree in applied physics from Universiti Putra Malaysia (UPM), in 2021. She has extensive research experience as a Data Science Research Assistant with the Department of Physics, UPM, where she focused on machine learning applications in medical diagnosis, including breast cancer detection and COVID-19 classification using radiomics features and automated machine learning techniques. She also has industry experience as a Data Analyst with Petronas Digital Sdn Bhd and Sprint Milestone Sdn Bhd. Her research interests include machine learning, data analytics, pattern recognition, medical imaging, automated machine learning, and drone imagery.

KHAIRUL ADIB YUSOF received the B.Sc. degree from Universiti Putra Malaysia (UPM), Seri Kembangan, Malaysia, in 2016, and the Ph.D. degree in space science from Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia, in 2021. He is currently a Senior Lecturer with the Department of Physics, UPM. His research interests include earthquake precursor, ground and space geomagnetic observations, signal processing, and computational physics.

ROMI FADILLAH RAHMAT (Member, IEEE) received the bachelor's and master's degrees from Universiti Sains Malaysia, and the Ph.D. degree from Universitas Sumatera Utara. He is a Researcher and a Lecturer with Universitas Sumatera Utara. Currently, he is the Vice Dean of the Faculty of Computer Science and Information Technology. His research interests include image processing, deep learning, brain-like computing, the Internet of Things, and mobile computing.