Cross-era Soccer Player Evaluation System based on Adaptive Weighting and Honor Competition Scheduling

Xingye Zhang
Computer and Artificial Intelligence Departments, Shandong University Finance and Economy, Jinan, Shandong, China

Abstract: Under the background of big data-driven sports analysis paradigm transformation, the comprehensive evaluation of soccer players faces the double challenges of insufficient cross-position fairness and static evaluation of honor value. Due to the solidified position weight and the neglect of the honor competition intensity in the traditional scoring system, the quantitative imbalance of the value of different role players and the lack of cross-era evaluation benchmarks are caused. This study proposes a comprehensive scoring model with cross-position adaptive weighting and honor dynamic weighting. By constructing an evaluation framework including attack, defense, organization, tactical role, and honor dimensions, the dynamic weight optimization is realized by combining Analytic Hierarchy Process (AHP) and random forest algorithm to solve the fair comparison problem under position differences. The technical scheme of the whole process from multi-source data collection, feature engineering to dynamic scoring was designed, and the "Honor competition intensity factor (CFI)" was innovatively introduced to calibrate the honor value based on the intensity of competition and the strength distribution of candidates. The system function module planning covers data processing, model training and visual analysis, and is expected to be applied to professional club recruitment decision-making, player market valuation and mass sports data analysis scenarios. This study provides a methodology innovation that integrates domain knowledge and data-driven for the field of intelligent sports analysis, and lays a technical foundation for cross-era player evaluation.

Keywords: Football Player Evaluation; Multi-Dimensional Weighted Scoring; Cross-Location Comparison; Honor Weight Adjustment; Machine Learning; Intelligent Sports Analysis

1. Introduction

As one of the most commercially valuable and socially influential sports in the world, the scientific evaluation of the comprehensive performance of football players has long been a key issue in sports industry decision-making, media content production, club management and fan culture construction. Although big data technology has been deeply involved in the field of sports analysis, the current mainstream scoring systems (such as FIFA Rating and WhoScored) still have structural defects, which make it difficult to accurately depict the multi-dimensional value of players. These platforms generally build evaluation models based on basic statistical data such as goal, assist and pass success rate, and their limitations are mainly reflected in two dimensions:

First, the fairness of cross-location evaluation is missing. The current system systematically undervalues defensive and organizational players. Taking the game of Barcelona against a team in the 11th round of La Liga in the 2024-25 season as an example, defenders Martinez and Cubarsi achieved zero seal defense through high-frequency clearance, interception and tactical offside (successfully restricting the offside of the opposing striker 8 times in a single game), but SofaScore only gave 7.3 points and 7.8 points respectively. Fellow forward Lewandowski scored 8.4 points with two goals. In fact, the tactical value and defensive contribution of the back line have the same weight as the offensive score in the decision of the game. However, the existing model relies too much on the offensive data, which leads to the quantitative imbalance of the value of players in different positions.

Second, the dynamic calibration mechanism of honor value is absent. The existing system adopts a static scoring mode for player honor, ignoring the difference in the intensity of competition for honor in different periods. Taking the Golden Ball as an example, in 2010, Messi, Iniesta and Xavi were the core of Barcelona and the Spanish national team when they were in their peak. Their personal data and team honors reached a historic high, and the difference in vote rate was only 6.17 percentage points, which was recognized as the "most competitive Golden Ball year". In 2022, Benzema has 549 points ahead of Salah (193 points), which is significantly lower than the intensity of competition in 2010. If a fixed score is given to the honors of different years, the World Cup champions in Maradona's era cannot be compared with the contemporary similar honors under the same value coordinate system, which seriously affects the scientification of cross-era evaluation.

In view of the above double challenges, this study constructs a cross-location adaptive weighting and honor dynamic weighting evaluation model, and achieves a breakthrough through the following innovations: ① A multi-perspective evaluation framework including attack, defense, organization, tactical role, honor and other dimensions was established, and the "Analytic Hierarchy Process (AHP) + random forest algorithm" hybrid modeling was used to realize the dynamic weight distribution of players in different positions. Secondly, the "Honor competition Intensity Factor (CFI)" is designed to calibrate the annual value of Ballon d 'Or, Champions League and other honors based on the intensity of competition and the strength distribution of candidates, so as to solve the problem of the benchmark unity of cross-era evaluation.

The research value is reflected in three aspects: ①Method innovation: breaking through the traditional static weight framework, establishing a position-sensitive dynamic weighting mechanism to provide technical solutions for cross-position player value comparability; ②Theoretical contribution: to construct the quantitative model of honor competition intensity, fill the theoretical gap of cross-era evaluation, and improve the methodology of sports statistics; ③Application expansion: Through periodic data update and multimodal visualization design, it provides standardized tools for professional clubs' recruitment decisions, media data narrative and fans' rational discussion, and promotes the transformation of football evaluation from empiricism to data science paradigm.

In summary, this paper focuses on the problems of "cross-position unfairness", "honor static weighted distortion" and "lack of dynamic feedback mechanism" in the current player scoring system, and proposes a multi-dimensional scoring framework and algorithm implementation scheme, which provides theoretical support and method foundation for the construction of future-oriented intelligent sports analysis system.

2. Related Research

The scientific modeling of soccer player's comprehensive scoring is the core proposition in the intersection field of sports statistics and data science. The existing research can be summarized into two types of technical paths: subjective experience oriented and data quantification oriented. The former is represented by the expert scoring system of FIFA, ESPN and other institutions, and relies on the professional judgment of industry practitioners. However, affected by factors such as media public opinion and individual cognitive bias, the cross-sample stability and repeatability of scoring results are insufficient (Wolf, 2021)[1]. The latter is subdivided into three streams of technology:

Subjective experience-based methods generate scores (e.g., FIFA award of the Year) through qualitative judgments of experts or media. Their value lies in the integration of unstructured field observations, but their significant shortcomings are the black-box scoring logic and the unreproducibility of results. The empirical study of Wolf (2021) shows that the difference in scoring standards of different experts on the same player in the same kind of competition can reach 15%, reflecting the structural problem of the lack of a unified quantitative benchmark in the subjective system[1].

The statistical weighted scoring model takes WhoScored and Sofascore as typical applications, and realizes quantitative evaluation by linear weighting of basic indicators such as shooting efficiency and passing success rate. Although these methods improve the objectivity of evaluation, there are three limitations: ①the weight system is preset based on the experience of domain experts, which can not be dynamically calibrated with the evolution of tactics; ② a unified index pool is used to cover all position players, ignoring the difference between the core ability of the forward and the defender (for example, the lack of quantification of the "offside number" of the defender); ③ Tactical contributions (such as the hub role in attack-defense conversion) are difficult to be effectively characterized by single-dimensional data (Pappalardo et al., 2019)[3].

The machine learning optimization model improves the weight allocation through the data-driven paradigm, such as the PlayeRank framework proposed by Pappalardo et al. (2019)[3], which optimizes the ranking of players in the same position through multi-dimensional feature vectors and supervised learning algorithms. However, its limitations are as follows: ① it only supports the horizontal comparison within the unit, and does not establish a unified benchmark for cross-role scoring; ② there is a lack of dynamic modeling of the honor data in the player's career, and it is difficult to answer the cross-era questions such as "How to quantitatively compare Maradona's 1986 World Cup with Messi's 2022 World Cup".

Graph modeling and complex network methods use graph convolutional neural networks (GCN) or centrality indicators (such as betweenness centrality) to quantify the tactical hub role of players (Chacoma, 2025) by constructing player passing networks and attack-defense linkage graphs[2]. However, these methods focus on the analysis of the dynamic structure of a single game, and lack the integration of the time dimension of the player's entire career data, which is difficult to meet the needs of long-term performance evaluation.

In the aspect of honor evaluation, the existing studies generally adopt a static scoring model (for example, the World Cup champion is assigned a fixed score of 20 points), ignoring the differences in competition intensity of the same honor in different years. For example, the gold content of the 2010 Champions League winner needs to be comprehensively evaluated by taking into account the ELO rating of the participating teams in that year and the injuries of key players, and static models cannot capture such dynamic characteristics. Although Jung et al. (2025) proposed the concept of competition difficulty coefficient based on ELO score, a complete modeling framework covering the intensity of honor competition has not yet been formed[5].

The limitations of the existing research are mainly reflected in the following aspects: ① insufficient position generalization ability: no unified scoring benchmark has been established across roles, which makes it difficult to compare the value of defenders and forwards under the same dimension; ② static honor evaluation: the lack of honor value calibration mechanism based on competition intensity and historical background, resulting in the lack of cross-era evaluation benchmarks; ③ weak system iteration: most models are static architectures with one-time training, which cannot respond to tactical evolution and data update requirements.

This study systematically solves the above problems by introducing a position-sensitive dynamic weighting mechanism, honor competition intensity factor (CFI) and a periodic model update framework, and provides a scientific and timely solution for player evaluation[4].

3. Method Design

3.1 System Architecture and Data Flow

The player evaluation system constructed in this study adopts a hierarchical technical architecture to automate the whole process from data collection to scoring output. Regarding data sources, the system primarily draws from three major platforms: Opta, WhoScored, and the official FIFA database. Together, these sources enable a comprehensive view of both in-game performance and historical accolades. Opta provides event-level data with spatial and temporal details for every pass, tackle, and interception, offering deep insights into player behavior and tactical execution. WhoScored supplements this with aggregated match ratings and technical statistics from top European leagues, such as shooting efficiency and passing accuracy. Meanwhile, the FIFA database includes structured records of individual honors like the Ballon d'Or and The Best FIFA Men's Player, as well as club and national team titles, along with basic career information. The integration of these sources ensures the evaluation model is built on a solid and multidimensional data foundation. The data layer regularly obtains multi-source data through customized crawlers, including technical statistics (such as goals, steals, offside induced times) and honorary records (such as Golden Ball, Champions League champions)[6]. Once collected, the raw data undergoes a thorough preprocessing phase before being stored and used. The first step involves format unification and duplicate removal to ensure consistency across sources. For missing values-such as incomplete match statistics-the system applies player-specific averages where appropriate, while non-critical or heavily missing entries are discarded. Outliers, like unusually high sprint distances or anomalous pass counts, are identified using statistical methods such as interquartile range and z-score analysis. Clear measurement errors are corrected when possible, and uncertain cases are flagged for manual review. Finally, inconsistent naming conventions, date formats, and categorical labels across different platforms are standardized to maintain coherence throughout the database. This data cleaning workflow ensures that the subsequent scoring process is both reliable and reproducible. After cleaning, the data are stored in relational database (MySQL) and unstructured database (MongoDB) respectively. The model layer combines domain knowledge and data-driven method: firstly, experts determine the initial weights of scoring dimensions through Analytic Hierarchy Process (AHP), and then the random forest algorithm is used to train the historical data to optimize the weights of each dimension and generate a comprehensive score. The back-end of the system provides interface services based on the Flask framework, and the front-end develops a visual interface through Vue.js to support player data query and multi-dimensional comparative analysis. The system architecture diagram is as follows in Figure 1:

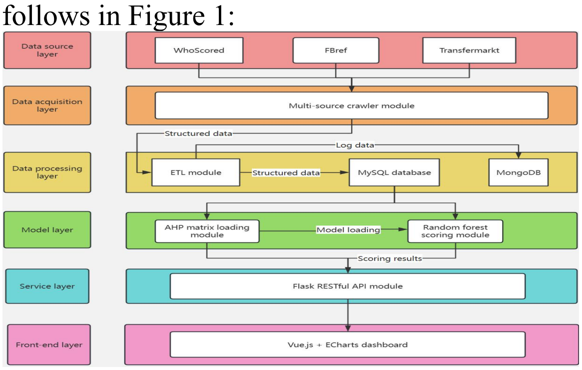

Figure 1. The System Architecture Diagram

3.2 Multi-Dimensional Scoring Index System

The scoring model comprehensively assesses player performance across six dimensions:

The attacking ability measures the direct attacking contribution by scoring goals, assists and shooting rate, and introduces advanced statistics such as "expected Goal (xG)" to evaluate the finishing efficiency.

In addition to the traditional interception and interception data, the defensive ability pays special attention to the "successful number of creating offside" and the success rate of defensive confrontation.

Through passing times, success rate and the number of opportunities created by passing, combined with the "centrality index" in passing network analysis, the organization ability evaluates the scheduling role of players in offensive and defensive transition.

The tactical value includes innovative indicators such as the number of successful downfield pressure and the distance of midfield return and defense to quantify the supporting effect of players on the tactical system.

Through the running distance, the number of sprints and the proportion of high-intensity running, the player's continuous combat ability was analyzed.

The static scoring model of honor achievement is abandoned, and the actual value of honor in different years is calibrated by dynamic factors[7].

3.3 Weight Optimization for Evaluations Across Locations

In order to solve the problem of fair evaluation of players in different positions, the system is modeled according to the classification of striker, midfielder, defender and goalkeeper:

Forwards focus on the weight of attack dimensions, such as goals, key passes and other indicators, accounting for more than 40%.

Defenders and goalkeepers strengthen the weight of defensive and tactical contributions, such as the number of successful offside creation, save success rate and other indicators accounted for no less than 35%;

Midfield players adopt a balanced weight strategy, and the proportion of organization, defense and physical dimensions is relatively average (about 25%-30%).

In the process of weight determination, firstly, the position-specific index hierarchy is constructed by AHP, and then the machine learning algorithm is used to train the data of each position, and the weight is dynamically adjusted to reflect the position characteristics and tactical evolution.

3.4 Dynamic Calibration Method for Honor Value

The Competitive Intensity Factor (CFI) is introduced in this study to quantify the contextual value of player honors based on the historical competitiveness of the award. Unlike static scoring systems that treat all honors equally regardless of era, the CFI dynamically adjusts honor scores by considering both the closeness of competition among candidates and the overall strength of participating teams.

The CFI is calculated using the following formula:

$$CFI_t = \alpha \times (\sigma_t / \sigma) + \beta \times \left(ELO_t / EL\overline{O}\right) \quad (1)$$

In this formula:

(1) σ_t is the standard deviation of vote shares or expert ratings among top award candidates in year t;

(2) $\bar{\sigma}$ is the historical average of σ_t over a reference period (e.g., the past ten years);

(3) ELO_t is the average ELO rating of the top teams relevant to the honor in that year;

(4) $EL\bar{O}$ is the historical average ELO rating.

The coefficients are set to $\alpha = 0.6$ and $\beta = 0.4$, reflecting the relative influence of rating dispersion and team strength. After calculating CFI_t, the original honor score H_t is adjusted using the formula:

$$H_t^{'} = H_t \times [1 + \gamma \times (CFI_t - 1)] \quad (2)$$

where $\gamma = 0.5$ is a smoothing factor that prevents excessive fluctuations due to small variations in competitiveness.

The full process includes five steps:

(1) Compute σ_t based on vote shares or expert ratings;

(2) Calculate ELO_t from the top relevant teams;

(3) Normalize both σ_t and ELO_t by dividing them by their respective historical averages;

(4) Apply the CFI formula to get CFI;

(5) Use the adjustment formula to revise H.

For example, in 2010, the Ballon d'Or race featured three tightly matched candidates-Messi, Iniesta, and Xavi-resulting in a high σ_t and strong ELO_t values, leading to an increased CFI and an upward score adjustment. Conversely, the 2022 edition, in which Benzema was a clear frontrunner, would produce a lower CFI and slightly reduced honor score. This mechanism ensures that identical awards earned in different historical contexts are fairly and dynamically evaluated[8].

3.5 System Development and Technical Implementation

The system adopts a front-end and back-end separation architecture: The back-end is developed using Python, data processing and model calculation are implemented based on the Flask framework, and structured and unstructured data are stored through MySQL and MongoDB. The front-end builds an interactive interface based on Vue.js, and uses ECharts to realize the visual display of the scoring results. At the deployment level, Docker containerization technology is used to improve portability, and Nginx is used to achieve load balancing. The model update mechanism supports periodic data training (e.g., weekly updates of season data), ensuring that the scoring system can respond to changes in player performance and tactical environment in real time.

4. Experimental Design and Expectations

This research designs a multi-dimensional experimental scheme around system functionality, algorithm effectiveness, evaluation fairness and dynamic adaptability. The specific content is as follows:

4.1 Functional Usability and User Experience Testing

Taking the front-end interaction of the system as the core, the reliability of player score query, multi-dimensional comparison and visual output functions is verified. The test scenario includes: after the user enters the player's name, the system needs to return a comprehensive score including attack, defense, honor and other dimensions within 2 seconds. It supports the comparative analysis of radar charts of 2-5 players across positions and seasons, and provides the report export function in PDF/Excel format. Through user questionnaire survey (N≥100) and log analysis, the expected interface friendliness score is ≥70%, and the response time of core operations is controlled within 2 seconds, which ensures the accuracy of data display and the fluency of interaction.

4.2 Algorithm Performance and Honor Evaluation Verification

The top ten Ballon d 'Or players from 2015 to 2024 are selected as verification samples, and their technical statistics (goals, assists, etc.) and honor data are input to generate a comprehensive score through the scoring system. Pearson correlation coefficient was used to test the correlation between the scoring results and the actual award ranking, and the expected correlation coefficient was ≥0.7, indicating that the system can effectively map the association between player performance and honor value. At the same time, the generalization ability of the model was evaluated by leave-one-out cross validation to ensure that the stability error of the scoring results under different sample subsets was ≤5%.

4.3 Cross-location Fairness and Weight Rationality Analysis

30 players each of striker, midfielder, defender and goalkeeper were randomly selected from the five major leagues to analyze the weight distribution of scoring dimensions in different positions. Key verification: whether the weight of attack dimension of forwards is significantly higher than that of other positions (expected ≥35%), whether the average score of defensive dimension of defenders is more than 20 points higher than that of forwards, and whether the weight of indicators such as the success rate of saves of goalkeepers accounts for ≥40%. The standard deviation of each dimension weight ≤0.2 was used as the standard to test the sensitivity and fairness of the scoring system to position differences, so as to avoid excessive dominance of offensive data on the overall score.

4.4 Dynamic Update and Competitive Intensity Calibration Test

Simulate the data update scenario during the season to verify the responsiveness of the system to real-time performance and reputation background: after loading the new game data every week, observe whether the player's rating reflects the current round's performance in time (e.g., a player scoring two goals in a single game increases his rating by 10-15%); Input the data of Ballon d 'Oreal candidates from different years (e.g., Messi in 2010 vs. Benzema in 2022), test whether the "Honor Competition Intensity Factor (CFI)" is correctly calculated (the CFI value in 2010 is expected to be 30-50% higher than that in 2022), and verify whether the honor points are dynamically adjusted with the CFI value. It is expected that the system can achieve weekly data-driven rating update, and the cross-decade honor calibration error is ≤8%.

4.5 Application Scenario Simulation Test

In view of the club's recruitment scenario, this paper selects the recruitment target data of a Premier League team in the past three years, uses the scoring system to generate a player value report, compares the actual recruitment effect (such as the number of appearances, key game contributions, etc.), and evaluates the system's ability to predict the player's potential. In the media report scenario, by simulating the visualization requirements of player comparison, it is verified whether the chart output by the system meets the professional and readability requirements of data narrative, and it is expected that the user's recognition of the analysis conclusion is ≥65%.

Through the above experiments, the scientificity of the system in the technical level, the practicability of the application level and the adaptability of cross scenarios are fully verified, which provides data support prospects for subsequent optimization.

5. Conclusion and Prospect

Aiming at the core defects of the traditional scoring system for soccer players, this study constructs a comprehensive evaluation system that integrates cross-position adaptive weighting and dynamic calibration of honor competition intensity. By introducing the hybrid modeling of analytic hierarchy process and machine learning, the fairness problem of player evaluation in different positions was solved. Through the design of "honor competition intensity factor" (CFI), the scientific quantification of cross-era honor value was realized. The system architecture covers the whole process of data acquisition, feature engineering, model training and visual analysis. Its innovation is reflected in the following aspects:

1. Methodology breakthrough: establish a position-sensitive dynamic weight mechanism, so that the value of different role players such as forwards and guards can be compared under the same dimension;
2. Theoretical contribution: propose an honor evaluation model based on competitive intensity, which fills the theoretical gap of cross-era player evaluation.
3. Application value: It provides a standardized tool for professional clubs' recruitment decisions, media data reports, and fans' rational discussions,

promoting the transformation of football evaluation from subjective experience to data science.

Although the system design has been completed in this study, the actual performance still needs to be further improved through data verification and model optimization. Future research directions include:

1. Technology deepening: introducing deep learning models such as Graph Neural Network (GNN) and Transformer to optimize the modeling accuracy of tactical role dimension and improve the quantification ability of complex team cooperation;
2. Data expansion: include players' injury history, training load and other physiological data to improve the assessment of physical ability dimensions; The social media public opinion data were integrated to analyze the additional effect of public influence on the value of players.
3. Scene extension: develop a real-time scoring module to support dynamic analysis of player performance in live matches; The scoring subsystem of club youth training players was constructed to assist the mining of potential new stars.
4. Cross-domain application: Extend this research method to basketball, football and other team sports, explore a universal athlete evaluation framework, and promote the universal development of intelligent sports analysis technology.

This study provides a theoretical innovation and practice-oriented solution for the field of sports data analysis, and its continuous optimization and application will help the digital transformation of the football industry and promote the deep integration of sports science research and technology application.

References

[1] Stephan Wolf, Maximilian Schmitt, Björn Schuller (2021). A football player rating system. Journal of Sports Analytics 6 (4), 243-257, 2021.

[2] Andres Chacoma (2025, February). Identification and Optimization of High-Performance Passing Networks in Football. arXiv preprint arXiv:2502.01444, 2025.

[3] Luca Pappalardo, Paolo Cintia, Paolo Ferragina, Emanuele Massucco, Dino Pedreschi, Fosca Giannotti (2019). PlayeRank: data-driven performance evaluation and player ranking in soccer via a machine learning approach. ACM Transactions on Intelligent Systems and Technology (TIST) 10 (5), 1-27, 2019.

[4] Fernández, J., & Bornn, L. (2018). Wide Open Spaces: A statistical technique for measuring space creation in professional soccer. Journal of Quantitative Analysis in Sports, 14(2), 71-80, 2018.

[5] Dong Hee Jung, & Jason J Jung. (2025). Data-driven understanding on soccer team tactics and ranking trends: Elo rating-based trends on European soccer leagues. PloS one, 20(2), e0318485, 2025.

[6] Routley, K., & Schulte, O. (2015). A Markov Game Model for Valuing Player Actions in Ice Hockey. Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence, 782-791, 2015.

[7] Dai Zhiwei. Application of digital physical training in the training of adolescent football players [J]. Sports, Games and Technology, 2025, (14):89-91.

[8] Peng Hongwei. Construction of a comprehensive quality evaluation model for junior high school football players driven by big data [J]. Sports, Culture and Technology, 2024(16):178-180.

