

# AI-Assisted Game Management Decisions: A Fuzzy Logic Approach to Real-Time Substitutions

Pedro Passos Farias<sup>1</sup>

<sup>1</sup>Institute of Computing – Universidade Federal Fluminense (UFF)

Av. Gal. Milton Tavares de Souza – 24210-346 – Niterói – RJ – Brazil

[pedropassos@id.uff.br](mailto:pedropassos@id.uff.br)

October 2025

## Abstract

In elite soccer, substitution decisions entail significant financial and sporting consequences yet remain heavily reliant on intuition or predictive models that merely mimic historical biases. This paper introduces a **Fuzzy Logic-based Decision Support System (DSS)** designed for real-time, prescriptive game management. Unlike traditional Machine Learning approaches that encounter a “predictive ceiling” by attempting to replicate human behavior, our system audits performance through an objective, rule-based inference engine. We propose a methodological advancement by reformulating the *PlayeRank* metric into a **Cumulative Mean with Role-Aware Normalization**, eliminating the play-time exposure bias inherent in cumulative sum models to enable accurate intra-match comparison. The system integrates this refined metric with physiological proxies (fatigue) and contextual variables (disciplinary risk modulated by tactical role) to calculate a dynamic **Substitution Priority ( $P_{final}$ )**. Validation via a case study of the 2018 FIFA World Cup match between Brazil and Belgium demonstrates the system’s ecological validity: it not only aligned with expert consensus on executed substitutions (e.g., Gabriel Jesus) but, crucially, identified high-risk scenarios ignored by human decision-makers. Specifically, the model flagged the “Fagner Paradox”—a maximum-priority defensive risk—minutes before a critical yellow card, and detected the “Lukaku Paradox,” where an isolated assist masked

a severe drop in participation. These results confirm that Fuzzy Logic offers a transparent, explainable, and superior alternative to “black-box” models for optimizing real-time tactical decisions.

## 1 Introduction

The financial stakes in modern soccer have never been higher, creating a vast economic disparity between clubs. In the 2023/24 season, Manchester City earned £175.9 million in prize money for winning the Premier League title ge (2025). In contrast, Fluminense, a Brazilian club with more modest economic power, received a total of US\$50.71 million for reaching the semifinals of the 2025 Club World Cup—a sum that could have escalated to US\$40 million for winning the finalESPN.com.br (s.d.).

Despite these immense financial consequences, one of the most impactful in-game decisions: player substitutions, often remains guided by intuition and limited visible cues. As top coaches like Guardiola acknowledge, the timing and choice of a substitution can be the difference between winning titles and missing crucial opportunities.

This gap is particularly evident because tactical decision-making during a soccer match is one of the most complex and high-impact processes in team management. Substitutions, changes in tactical posture, or defensive adjustments are frequently made by coaches based on a combination of empirical knowledge and a multifactorial assessment of inherently imprecise variables. Factors like ”the player looks tired,” ”performance has recently dropped,” or ”the risk of expulsion is high” are not binary, but rather gradual. Performance evaluation is, in itself, a notable challenge. Although the availability of large-scale event data has grown, there is no single, universally accepted metric that captures all facets of a player’s performance Pappalardo et al. (2019).

This work addresses the problem of tactical decision-making under uncertainty through the implementation of a fuzzy logic control system. Fuzzy logic is an artificial intelligence tool particularly suited for this domain, as it allows for modeling ”expert knowledge” (a coach’s rules) and handling vague linguistic concepts, such as ”high age,” ”low performance,” or ”medium fatigue” (Marliere, 2017). The objective of this paper is to design,

implement, and validate a decision support system (DSS) focused specifically on player substitution priority. The system uses four inputs (performance, fatigue, age, and disciplinary risk) to generate a fuzzy output (Substitution Priority), providing a quantitative tool to assist the coaching staff with one of their most critical real-time decisions.

## 2 Literature Review

The academic literature on soccer analytics provides the theoretical foundation for this work, situated at the intersection of performance evaluation and tactical decision modeling. A central methodological challenge in this domain is quantifying player performance. The **PlayeRank** framework, proposed by Pappalardo et al. (2019), defines a multidimensional, role-aware metric that combines individual technical actions with influence on the passing network, validated against professional scout assessments. The authors explicitly aim to reduce biases related to playing time, normalizing performance vectors by the number of actions to “avoid biases due to play time.”

However, although normalization occurs at the level of the action vector, the final match score is computed as a weighted cumulative sum of contributions. When inspected over the temporal progression of a match, this formulation tends to produce strictly increasing trajectories, effectively reintroducing a play-time exposure bias that favors players who remain longer on the field, even if their per-minute contribution rates are similar. This monotonic behavior introduces a methodological limitation for temporal or comparative analyses. Each additional action contributes positively to the cumulative sum—regardless of qualitative impact—so players with identical per-minute performance profiles diverge in total score purely as a function of playing time. While this approach is suitable for match-level summaries, it becomes problematic when the goal is to characterize relative influence or detect shifts in performance dynamics.

The strictly increasing nature of the cumulative score also prevents the representation of performance decay, momentum reversals, or tactical suppression: a player whose influence declines due to fatigue or tactical adjustments will still exhibit an artificially monotonic performance trajectory. Therefore, while PlayeRank’s cumulative formulation is adequate for full-match evaluation, it is insufficient for temporal segmentation. A temporally meaningful metric must reflect evolving game states, substitutions, and local contributions, rather than indiscriminately accumulating actions.

Recent work supports this perspective. Schmidt et al. (2024) demonstrate that player influence fluctuates meaningfully over a match due to tactical reconfiguration, opponent

pressure, and physical factors, advocating for temporal windowing to capture such dynamics. Following this insight, the present work adopts 5-minute temporal slices, but advances the methodology by replacing the cumulative sum with a cumulative mean across slices. This modification preserves interpretability while enabling direct comparability between players with different playing times, eliminating residual exposure bias and allowing the performance curve to reflect true temporal variability.

A second methodological challenge concerns the interpretation of imprecise or uncertain information in decision-making contexts. The literature highlights fuzzy logic as an effective tool for this purpose. For instance, Huarachi-Macuri et al. LACCEI (2023) demonstrate the suitability of fuzzy systems for transforming quantitative descriptors into interpretable linguistic assessments (e.g., Stamina, Agility). Within tactical modeling, Marliere (2017) uses fuzzy control systems to arbitrate between tactical states under uncertainty. These contributions motivate the role of fuzzy logic in the present system: integrating performance indicators and contextual factors into a Substitution Priority score in a manner compatible with the inherent uncertainty of in-game decision making.

The decision to make substitutions in soccer remains an intrinsically human domain, hardly replicable by machine learning (ML) algorithms. While ML can analyze historical performance data, a coach’s choice involves a complex assessment of contextual, tactical, and psychological factors that are not easily quantifiable—such as team morale, perceived fatigue, reading the opponent’s strategy, and game momentum. Attempts to predict substitutions using ML, while informative, reveal a clear ceiling in their ability to capture this complexity. A recent study by Mohandas et al. Mohandas et al. (2023), for example, analyzed a large dataset of 51,738 substitutions using multiple algorithms, including Random Forest, SVM, and Decision Trees. Even with this wealth of data, the best-performing model (Random Forest) achieved a maximum accuracy of just over 70%. This gap of approximately 30% is significant, suggesting that the final decision is not purely predictive but rather adaptive, relying on subjective human judgment that historical data alone cannot capture. Consequently, models attempting to predict the optimal substitution are conceptually limited. Alternative approaches, such as fuzzy logic-based decision support

systems, appear more suitable, as they are designed not to replace but to support the coach’s judgment, managing the inherent uncertainty and qualitative factors of the match.

Despite these advances, a gap remains in the literature. Existing works either focus on performance evaluation (Pappalardo et al., 2019; LACCEI, 2023) or on tactical reasoning frameworks Marliere (2017), but there is no decision support system that integrates a validated performance metric with contextual match factors—such as fatigue, momentum shifts, and disciplinary risk—for the specific purpose of informing substitution decisions. This work addresses this gap by combining temporally resolved performance evaluation with fuzzy decision modeling within a unified substitution strategy framework.

### 3 Dataset

The database selected for this project is the “*Soccer match event dataset*”, a detailed public repository of soccer match events (Pappalardo, 2020). The choice of this dataset is based on its high granularity. The *dataset* details player-level actions and aggregated performance metrics, allowing for the modeling of *in-game* situations, which is an essential requirement for our system. We used the complete *dataset* available on Kaggle (Pappalardo, 2020), which comprises 27 interrelated CSV tables. They cover seven main competitions and can be grouped into five logical categories:

- **Match Data (`matches_*.csv`):** Contains information about the games, such as dates, lineups, substitutions, results, and tactical formations.
- **Event Data (`events_*.csv`):** The core of the *dataset*, recording millions of individual actions on the field.
- **Entity Data (`players.csv`, `teams.csv`, etc.):** Dimensional tables with demographic and static data.
- **Performance Metrics (`playerank.csv`):** A pre-processed file that provides the `playerankScore`, a multidimensional and role-aware performance evaluation metric.

- **Dictionaries (`tags2name.csv`, `eventid2name.csv`):** Metadata that translate event and *tag* IDs into readable descriptions (e.g., Tag 1702 = 'yellow\_card').

The choice of `playerankScore` as our main "Performance" input is a central methodological decision. As proposed by Pappalardo et al. (2019), the PlayeRank *framework* was developed to solve the absence of a consolidated and universally accepted metric for evaluating player performance. The `playerankScore` is a metric derived from millions of game events that, according to the authors, surpasses other metrics when compared with assessments from professional scouts. Therefore, instead of trying to model performance from raw events, we adopted the `playerankScore` as an already validated and academically robust representation of a player's performance in a match.

Dataset available at:

```
https://www.kaggle.com/datasets/aleespinosa/soccer-match-event-dataset/data?  
select=playerank.csv
```

## 4 Data Integration and Pre-processing

The pre-processing pipeline transforms heterogeneous soccer event logs into a unified temporal representation of player performance suitable for fuzzy inference. The process unfolds in three sequential stages: temporal performance computation, contextual and demographic enrichment, and role-aware normalization. Each stage progressively increases the dataset's semantic density and interpretability.

The objective of this pipeline is to transform the raw multi-source event data (events, matches, players) into a single structured and temporal dataset tailored for the Fuzzy Logic Decision System. Unlike the original *PlayeRank* framework Pappalardo et al. (2019), which aggregates player performance per match, our system requires a fine-grained temporal perspective of the player's performance evolution within a match. To achieve this, we implemented a three-stage processing pipeline that generates a dense temporal dataset, where each record represents the state of a player in a 5-minute (300-second) interval of

play.

## 4.1 Stage 1: Temporal Performance Metric Generation

The first stage (`playerankdatasetV4.py`) processes the raw event logs (`events_*.csv`) and converts each action’s timestamp (`eventSec`, `matchPeriod`) into an absolute measure of seconds from the start of the match (`total_seconds`). The match timeline is then discretized into 5-minute intervals, associating each event with its corresponding temporal slice. The following variables are created in this stage:

- `matchId`, `teamId`, `playerId`: extracted directly from the event logs, providing the hierarchical identifiers for each observation.
- `cartao_amarelo`: a binary state indicator (using `expanding().max()`) equal to 1 if the player received a yellow card at any point up to the current slice.

Within each interval, two complementary components are calculated:

- **Technical Score (`score_tecnico_fatia`)**: the aggregation of event-based actions (passes, shots, duels, interceptions) using the same weights of the original PlayeRank framework. Events are mapped to technical dimensions and normalized within the team context.
- **Network Score (`score_rede_fatia`)**: computed using a directed pass graph per team and interval. Following the original PlayeRank framework, the *Eigenvector Centrality* of each player is used to quantify their structural influence on ball circulation.

These two components are linearly combined to form the primary performance indicator for each temporal slice:

$$\text{playerank\_fatia\_raw} = (1 - \alpha_{\text{net}}) \cdot \text{score\_tecnico\_fatia} + \alpha_{\text{net}} \cdot \text{score\_rede\_fatia}$$

where  $\alpha_{\text{net}} = 0.2$  controls the relative contribution of the network-based component. This configuration was selected to balance structural influence and technical efficiency, aligning with the sensitivity ranges reported in Pappalardo et al. (2019).

**Cumulative Metric Redefinition.** To capture performance evolution through time, we define the cumulative mean score:

$$\text{playerrank\_acumulativo\_media\_raw} = \frac{1}{t} \sum_{i=1}^t \text{playerrank\_fatia\_raw}^{(i)}$$

This cumulative formulation departs from the additive model of the original PlayeRank, ensuring that longer playing times do not artificially inflate the total performance. The metric represents the player's average contribution up to time  $t$ , maintaining comparability among players with different play durations.

## 4.2 Stage 2: Contextual and Demographic Enrichment

The second stage (`expandirdataset.py`) enhances the temporal dataset with contextual, situational, and demographic variables by joining multiple data sources.

- **player\_age**: calculated by merging player metadata (`players.csv`) with match dates and computing the player's age on match day.
- **player\_position**: extracted from the player metadata, representing the tactical role (goalkeeper, defender, midfielder, forward).
- **goals\_scored**: identified from events where `eventName = "Shot"` and `subEventName = "Goal"` and accumulated over time.
- **assists**: extracted using Wyscout tag `id = 302` and accumulated over time using `groupby(['matchId', 'playerId']).cumsum()`.
- **momentum\_rate**: a measure of short-term performance trend, calculated as the difference between the current( $t$ ) and the previous( $t - 1$ ) 5-minute `playerrank_acumulativo_media_percentil` slices.

This enrichment process yields a semantically rich dataset combining technical, contextual, and demographic aspects for each player's temporal trace.

### 4.3 Stage 3: Final Cleaning and Role-Aware Normalization

The final stage (`limpezafinal.py`) standardizes the temporal data and recalculates cumulative metrics for stability and accuracy. The following operations are performed:

- **removal of out-of-field periods:** Periods where it was identified that the player was substituted, or entered as a substitute, were removed to ensure that the metric remained accurate, not applied to goalkeepers
- **Temporal Harmonization:** The temporal variables are standardized. `Tempo_Partida` is set to represent the \*end\* of each 5-minute slice (e.g., 5, 10, 15...), and `minutes_played` is recalculated using `cumcount()` to ensure a precise cumulative sum of on-field time.
- **Recalculation of Cumulative Means:** The variable `playerank_acumulativo_media_raw` is recalculated using a robust `expanding().mean()` to ensure that averages reflect the final, harmonized temporal trace.
- **Percentile Normalization:** two percentile-based variables are computed within each (`matchId`, `position`) group. Note: this normalization uses the raw `position` column (from the original data source) to group players by tactical role, ensuring methodological consistency:

$$\text{playerank\_fatia\_percentil} = \text{rank}_{\text{pct}}(\text{playerank\_fatia\_raw})$$

$$\text{playerank\_acumulativo\_media\_percentil} = \text{rank}_{\text{pct}}(\text{playerank\_acumulativo\_media\_raw})$$

This *role-aware normalization* ensures that performance is assessed relative to tactical peers, preserving fairness and interpretability in subsequent fuzzy inference.

**Final Dataset Variables.** The resulting dataset (`Dataset_limpo_final.csv`) contains the following variables:

- **Identifiers:** matchId, playerId, teamId.
- **Temporal Dimensions:** Tempo\_Partida, minutes\_played.
- **Performance Metrics:** playerank\_fatia\_raw, playerank\_acumulativo\_media\_raw, score\_tecnico\_fatia, score\_rede\_fatia.
- **Normalized Metrics:** playerank\_fatia\_percentil, playerank\_acumulativo\_media\_percentil.
- **Contextual Factors:** momentum\_rate, cartao\_amarelo.
- **Demographics:** player\_age, player\_position.
- **Offensive Contributions:** goals\_scored, assists.
- **Auxiliary:** position, a placeholder variable maintained for compatibility.

Each record corresponds to a 5-minute temporal snapshot of an individual player's on-field performance, integrating technical, contextual, and demographic information in a consistent temporal framework.

#### 4.4 Data Validation

Before integration into the fuzzy system, consistency checks were performed. The recalculation of `playerank_acumulativo_media_raw` in the final stage (Stage 3) ensures that all cumulative metrics are based on the robust `expanding().mean()` operator applied to the complete and harmonized temporal trace of each player. This guarantees that all cumulative scores represent valid in-game performance dynamics and that the fuzzy system operates on stable, correctly aggregated data.

### 5 Exploratory Data Analysis

The Exploratory Data Analysis (EDA) was conducted on the final integrated dataset, consisting of 805,146 temporal observations, 3,035 unique players, and 1,941 matches.

The analysis aimed to (i) characterize the statistical properties of the input variables used in the fuzzy inference system, (ii) validate soccer-specific behavioral hypotheses (e.g., positional risk asymmetry), and (iii) provide empirical grounding for the design of the fuzzy universes of discourse.

## 5.1 Descriptive Overview

The dataset preserves a balanced representation of real-world soccer demographics and tactical structures. Table 1 summarizes the player distribution and key descriptive statistics.

Table 1: Dataset Summary and Descriptive Statistics

Category	Value	Metric	Notes
Total observations	805,146	intervals	5-minute resolution
Unique players	3,035	–	across 7 competitions
Matches analyzed	1,941	–	complete event coverage
Mean/Median performance (playerank_acumulativo_media_percentil)	0.501 0.501	std = 0.289 –	centered around 0.5 symmetrical distribution
Player positions	4 roles	–	role-aware normalization applied
Goalkeeper	233	(5.8%)	
Defender	1,032	(25.8%)	
Midfielder	1,120	(28.0%)	
Forward	650	(16.3%)	

This composition closely mirrors elite-level soccer distributions, where midfielders and defenders together comprise over half of the active roster, reflecting the denser tactical occupation of central zones.

## 5.2 Univariate Distributions

**Player Age.** The age distribution (Figure 1) follows an approximately median centered around 26.0 years, with a range between 15.0 and 45.0 years. The kurtosis (2.91) and near-zero skewness confirm symmetry. This validates the fuzzy partitioning into *Young*, *Peak*, and *Veteran* sets, with smooth transitions around the mean.

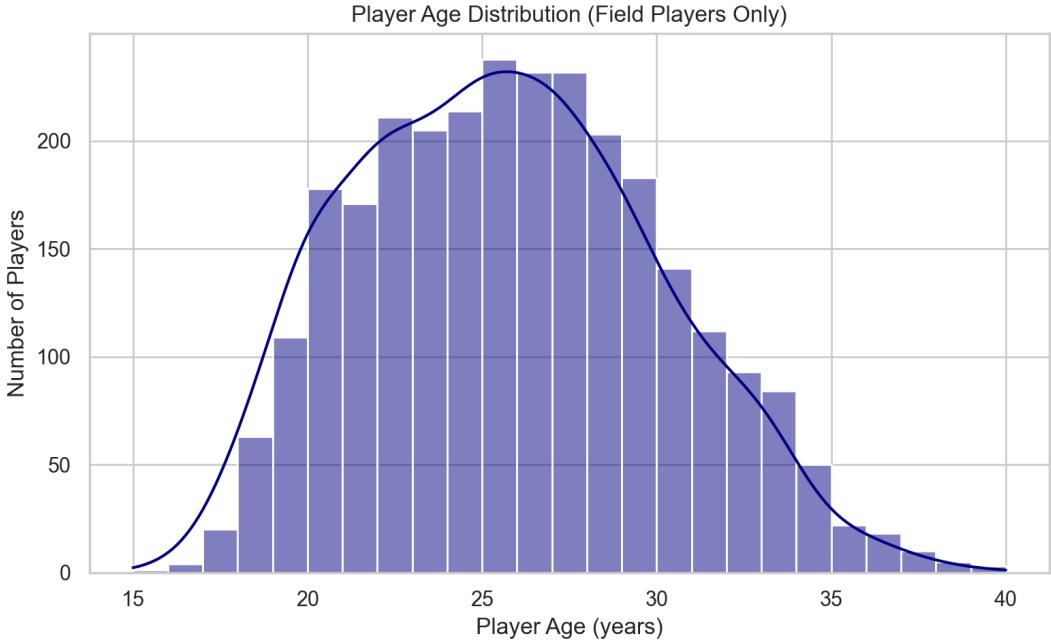


Figure 1: Distribution of player ages in the dataset.

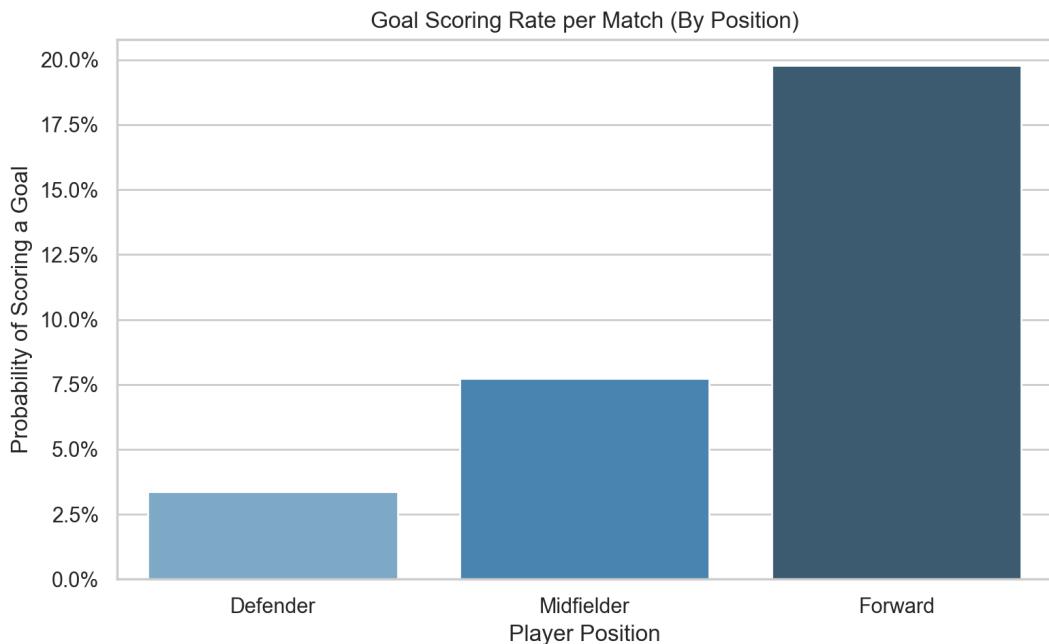
**Performance (`playerank_acumulativo_media_percentil`)** The per-slice performance scores exhibit a bounded, quasi-normal distribution centered around 0.5, with a standard deviation of 0.289. These characteristics make the metric well-suited for fuzzy modeling, where zero-centered or bounded scales allow intuitive linguistic partitioning (*Low*, *Medium*, *High*) without rescaling distortions.

### 5.3 Role-Based Event Distributions

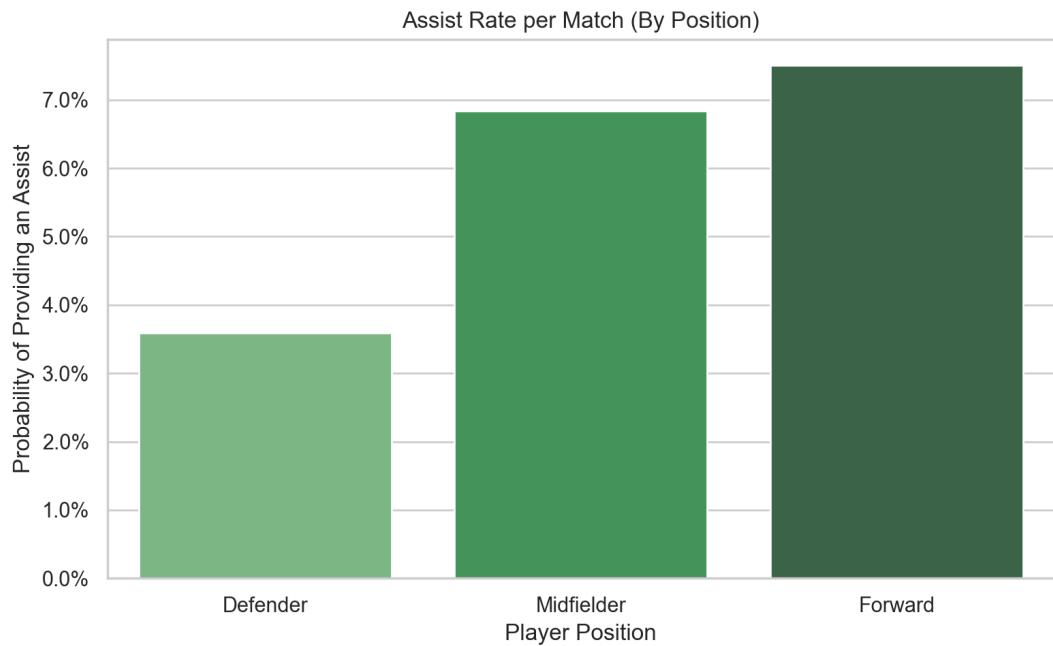
To further interpret positional behavior, the dataset was aggregated by player position to estimate empirical probabilities of key in-game events (goals, assists, and yellow cards).

The empirical distributions align with established tactical expectations:

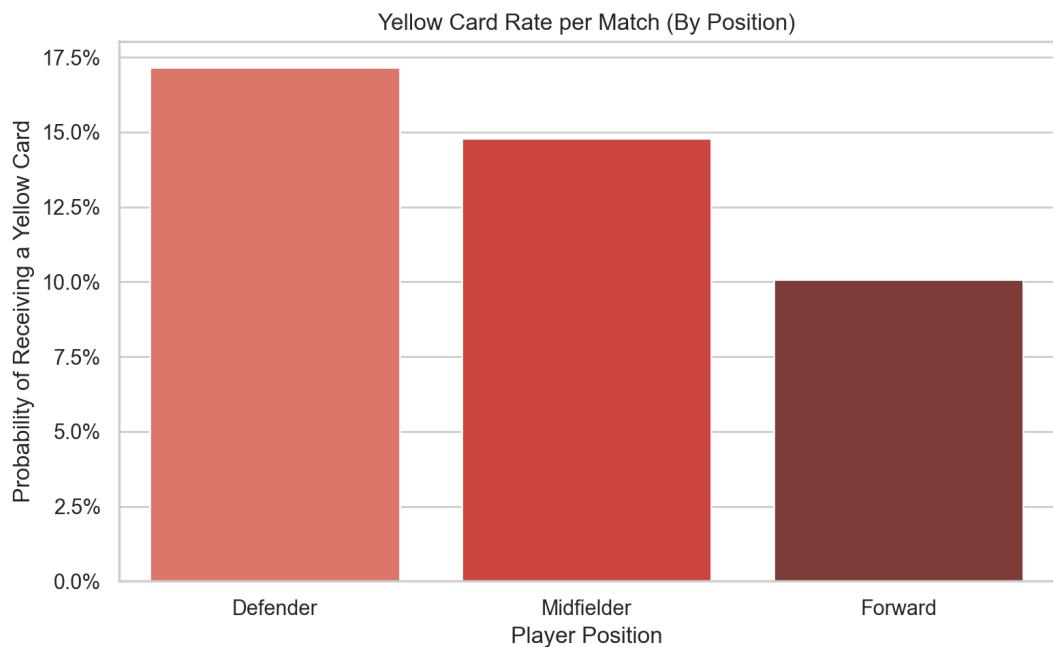
- **Goal Rate:** Forwards lead with a 19.8% probability of scoring per match, compared to 7.7% for midfielders and 3.4% for defenders. Goalkeepers are near zero, validating the offensive gradient embedded in the *player\_position* variable.



- **Assist Rate:** Forwards exhibit the highest assist probability (7.5%), reflecting their creative and distributive role in ball progression. Midfielders follow closely (6.8%), while defenders contribute the least(3.6%).



- **Disciplinary Risk:** Defenders show the highest yellow card rate (17.2%), followed by midfielders (14.8%) and forwards (10.1%), reinforcing the need for *contextual fuzzy rules* that modulate card risk according to tactical role.



## 5.4 Variable Definition, Independence and Correlations

### 5.4.1 Variable Definition and Selection

From the final integrated dataset (`Dataset_limpo_final.csv`), nine variables were selected for exploratory analysis and subsequent use as inputs in the fuzzy decision model. These variables represent the main dimensions of in-game player state: performance, fatigue, disciplinary risk, offensive contribution, and tactical context. Table 2 summarizes the selected features, their measurement scale, and their conceptual role within the system.

Table 2: Variables selected for exploratory analysis and fuzzy modeling.

Dimension	Variable (ID)	Mathematical Description and Relevance Set
Technical Performance	P_cum	$[0, 1] \subset \mathbb{R}$
Fatigue	Min_played	$[0, 100] \subset \mathbb{R}$
Disciplinary Risk	Card_Y	$\{0, 1\} \subset \mathbb{Z}$
Form Trend	Momentum	$[-1, 1] \subset \mathbb{R}$
Contextual Fatigue (Age)	Age	$[15, 45] \subset \mathbb{N}$
Tactical Role	Position	Categorical
Offensive Contribution (Goals)	Goals	$\mathbb{N}$
Offensive Contribution (Assists)	Assists	$\mathbb{N}$

**Labels:** P\_cum = playerank\_acumulativo\_media\_percentil, Min\_played = minutes\_played, Card\_Y = cartao\_amarelo, Momentum = momentum\_rate, Age = player\_age, Position = player\_position, Goals = goals\_scored, Assists = assists,

These variables were derived directly from the processed PlayeRank-based dataset, ensuring temporal coherence across 5-minute intervals. The combination of continuous, discrete, and categorical features provides a multidimensional view of player state suitable for fuzzy inference. In the next subsection, we assess the independence and correlation

among these variables to confirm their suitability for use as fuzzy inputs.

#### 5.4.2 Spearman Correlation Matrix

Figure 2 reports the Spearman correlation matrix computed over all input variables adopted in the fuzzy inference system. The pairwise correlations are consistently weak, with magnitudes predominantly below  $|\rho| < 0.20$ . Such a pattern indicates that the variables do not exhibit meaningful monotonic dependence.

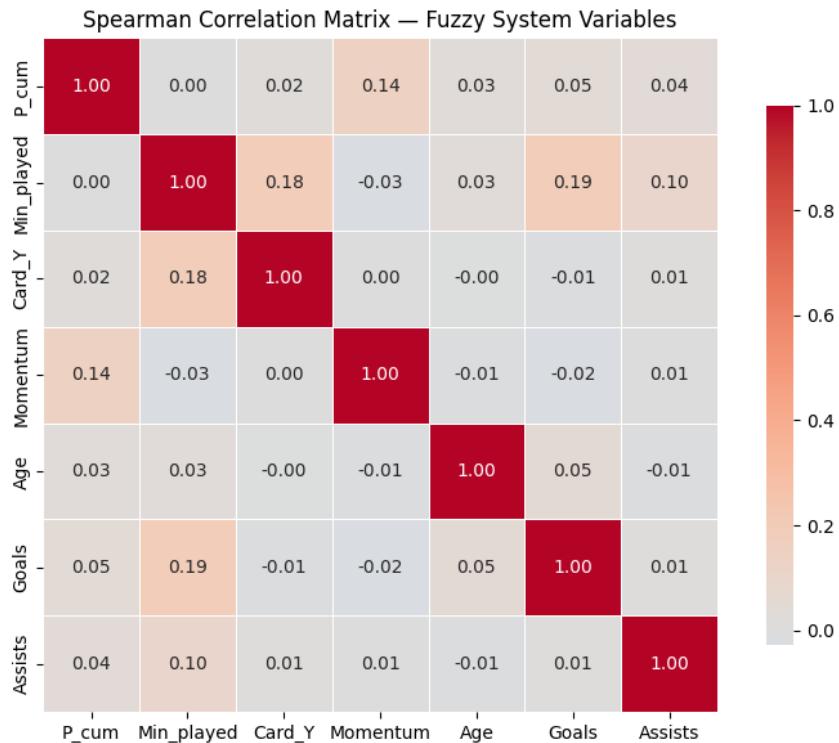


Figure 2: Spearman correlation matrix among the fuzzy-system input variables. The weak correlations indicate low redundancy and support the assumption of variable independence in fuzzy inference.

This low degree of association is desirable in the context of fuzzy inference. Since the rule base relies on linguistic terms and membership functions that encode distinct semantic dimensions of player performance, strongly correlated variables would introduce redundancy and reduce the discriminative power of the fuzzy rules. Conversely, the weak correlations observed here suggest that each variable contributes independent information to the inference process, thereby supporting a more expressive and interpretable fuzzy model.

## 6 Fuzzy Control System

Based on the theoretical foundation and the need for dynamic real-time evaluation, we designed the Fuzzy Control System (FCS) to function as a *Contextual Modifier*. Unlike traditional systems that calculate a raw output from zero, this architecture calculates a correction factor applied to a baseline performance metric. The system was implemented using the `scikit-fuzzy` library in Python, utilizing the Mamdani architecture.

### 6.1 Definition of Fuzzy Variables

The system architecture has been expanded to process high-dimensional match data. It comprises eight input variables (antecedents) and one output variable (consequent). The universes of discourse were calibrated based on the range of values observed in the dataset and domain constraints (e.g., match duration).

#### 6.1.1 Antecedents (Inputs)

1. **Cumulative Performance (P\_cum):** Represents the player's average percentile performance up to the current moment.
  - Universe: [0.0, 1.0].
  - Sets: *VeryLow, Low, Medium, High, VeryHigh*.
2. **Momentum (Momentum):** Rate of change in performance, indicating if the player is improving or declining.

- Universe: [-1.0, 1.0].
- Sets: *Falling, Stable, Rising.*

3. **Fatigue (Min\_played):** Minutes played in the match.

- Universe: [0, 100] minutes.
- Sets: *Low (0-45'), Medium (40-80'), High (70-100').*

4. **Age (Age):** Player's chronological age.

- Universe: [15, 45] years.
- Sets: *Young, Peak, Veteran.*

5. **Match Events (Categorical/Integer):**

- **Card\_Y:** Yellow card status [0, 1] (*Yes*).
- **Goals:** Goals scored [0, 10] (*None, Some, Many*).
- **Assists:** Assists provided [0, 10] (*None, Some, Many*).

6. **Positional Context (Binary):**

- Variables: `is_Defender`, `is_Midfielder`, `is_Forward`.
- Sets: *Yes* (indicating the player's role to activate specific rules).

### 6.1.2 Consequent (Output)

- **Modifier Value (Modifier\_Value):** A correction factor ranging from negative (protection/priority reduction) to positive (urgency/priority increase).
  - Universe: [-100, 100].
  - Granularity: 9 sets ranging from *VLN* (Very Large Negative, -100) to *LP\_70* (Large Positive, +70).
  - *Zero* represents no adjustment.

## 6.2 Membership Functions

The membership functions (MFs) mix Trapezoidal (`trapmf`) and Triangular (`trimf`) shapes to capture specific tactical thresholds. Key definitions from the implementation include:

- **P\_cum:** *Low* uses a trapezoid [0, 0, 0.10, 0.35] to capture distinctively poor performance, while *High* starts at 0.65.
- **Momentum:** *Falling* is defined strictly in the negative range [-1.0, -1.0, -0.03, -0.01], ensuring only genuine performance drops trigger the logic.
- **Fatigue:** Overlapping sets allow smooth transitions. *High* fatigue begins notably at 70 minutes, aligning with common substitution windows.
- **Contextual Events:** Variables like `is_Defender` or `Card_Y` use pseudo-binary trapezoids (e.g., [0.5, 1, 1, 1.5]) to function as logical switches within the fuzzy inference engine.

## 6.3 Rule Base

The rule base was streamlined to remove noise and focus on Intensifying or attenuating the player's performance using other contextual events as variables. The system avoids oscillation by ignoring minor fluctuations and focuses on critical states. The logic integrates position and stats directly. :

Table 3: Key Fuzzy Rules for Substitution Priority

ID	Rule Logic
<b>R01 – Untouchable Star</b>	IF (P_cum is High/VeryHigh) THEN (Modifier = Negative).
<b>R02a – Critical Fatigue</b>	IF (P_cum Low/VeryLow) AND (Min High) THEN (Modifier = Large Positive).
<b>R02b – Early Fatigue</b>	IF (P_cum Low/VeryLow) AND (Min Medium) THEN (Modifier = Medium Positive).
<b>R03 – Defensive Risk</b>	IF (Defender AND YellowCard Yes) THEN (Modifier = Medium Positive).
<b>R04 – Rapid Decline</b>	IF (P_cum Low/VeryLow AND Momentum Falling) THEN (Modifier = Large Positive).
<b>R07 – Positive Momentum</b>	IF (P_cum High/VeryHigh AND Momentum Rising) THEN (Modifier = Medium Negative).
<b>R08 – Ineffective Forward</b>	IF (Forward AND P_cum Low AND Goals None) THEN (Modifier = Large Positive).
<b>R09 – Striker Under Pressure</b>	(1) IF (Forward AND P_cum Low/Med AND Momentum Falling) THEN (Modifier = VeryLarge Positive). (2) IF (Forward AND P_cum Low/Med AND Min High) THEN (Modifier = VeryLarge Positive).
<b>R10 – Invisible Playmaker</b>	IF (Midfielder AND P_cum Low AND Assists None) THEN (Modifier = Med-Large Positive).
<b>R11 – Creator Bonus</b>	IF (Assists Some OR Goals Some) THEN (Modifier = Medium Negative). IF (Assists Many OR Goals Many) THEN (Modifier = Large Negative).
<b>R12 – Veteran Fatigue</b>	IF (Age Veteran AND Min High) THEN (Modifier = Medium Positive).
<b>R13 – Young Talent Protection</b>	IF (Age Young AND (Goals Some/Many OR Assists Some/Many)) THEN (Modifier = VeryLarge Negative).
<b>R14 – Goal Protection</b>	IF (Goals Some) THEN (Modifier = Medium Negative). IF (Goals Many) THEN (Modifier = Large Negative).
<b>R15 – Neutral State</b>	IF (P_cum Medium AND Momentum Stable) THEN (Modifier = Zero).

## 6.4 Inference and Final Calculation

The system employs a hybrid calculation model. The Fuzzy Inference System (FIS) computes the `Modifier_Value` using the Centroid method. However, this value is not the final priority. The final substitution priority  $P_{final}$  is calculated by applying the modifier to a baseline inverse of the cumulative performance, scaled by a factor  $\alpha$ :

$$Baseline = 100 \times (1.0 - P_{cum}) \quad (1)$$

$$P_{final} = \text{clip} (\text{Baseline} + (\text{Modifier} \times \alpha), 0, 100) \quad (2)$$

Where  $\alpha = 0.25$  in the current version. This ensures that the fuzzy logic acts as a "tuner" that intensifies or attenuates the urgency based on tactical context (e.g., a yellow card or a goal scored), rather than overriding the player's core performance metric entirely. This methodology allows the expert's knowledge ESPN Brasil (s.d.); Terra Esportes (s.d.); CNN Brasil Esportes (s.d.) to be faithfully captured by the system, adding significant weight to the player's performance on the field.

## 7 Results and Discussion

Tactical Analysis - Team 6380 (Match 2058011)



Figure 3: (each point represents the interval up to the next, e.g., 0 = 0–5 min, 5 = 5–10 min, 10 = 10–15 min)

To assess the ecological validity and practical utility of the Fuzzy Control System (FCS) in a high-stakes environment, the model was applied to historical data from the 2018 FIFA World Cup quarter-final match between Brazil and Belgium. This case study was selected due to the critical nature of the tactical decisions required under pressure

(with Brazil trailing 0-2 at halftime) and the availability of extensive post-match expert analysis to serve as ground truth for validation.

## 7.1 Adherence to Technical Commission Decisions

The simulation demonstrated a significant correlation between the *Substitution Priority* ( $P_{final}$ ) generated by the model and the actual alterations made by the coaching staff. This alignment suggests that the fuzzy inference engine successfully identified performance bottlenecks in real-time. Table 4 summarizes the critical decision windows, comparing the model’s output—specifically the ranking of the player within the team at that specific minute—with the real-world outcome.

Table 4: Comparison between Model Priority ( $P_{final}$ ) and Real Tactical Decisions (Brazil vs. Belgium, 2018).

Player	Analysis Slice( $t$ )	$P_{final}$ in the slice	Rank at $t$	Predominant Fuzzy Logic	Real Decision
Willian	40'-45'	72.0	4 <sup>th</sup>	Rapid Drop (R04)	Substituted out
Gabriel Jesus	55'-60'	99.1	2 <sup>nd</sup>	Ineffective Forward (R08)	Substituted out
Paulinho	60'-65'	93.1	2 <sup>nd</sup>	Ineffective Midfielder	Substituted out
Fagner	45' – 85'	100.0	1 <sup>st</sup>	Fatigue	Not Substituted

### 7.1.1 Analysis of Executed Substitutions

The model correctly flagged the three players substituted by the head coach, Tite, as high-priority candidates for removal, validating the specific fuzzy rules designed for performance decay:

- **Willian (45')**: At halftime, the system assigned a priority of  $P_{final} = 72.0$ . While he was ranked 4<sup>th</sup> in substitution urgency, the system triggered the *Rapid Drop* rule (R04), identifying a sharp decline in technical actions during the final 20 minutes of the first half. The coach’s decision to substitute him immediately supports the model’s diagnosis of offensive ineffectiveness, later described by media analysis as a performance “below expected levels” ESPN Brasil (2018).
- **Gabriel Jesus (58')**: Approaching the hour mark, the model calculated a critical priority of **99.1**, placing him as the 2<sup>nd</sup> highest priority in the squad. The system

activated rule R08 (*Ineffective Forward*), triggered by the combination of a low cumulative performance score ( $P_{cum}$ ) and a null value for goals and assists. Post-match statistics revealed the player had only 12 ball touches in 58 minutes Revista Época (2018), empirically validating the fuzzy system’s ability to detect tactical isolation.

- **Paulinho (73’):** With a priority of **93.1** till the minute 65 (last on-field activity registered), the model identified ”Early Fatigue” combined with low structural participation. This aligns with external critiques noting the player ”did not make his presence felt” and failed to provide the necessary offensive surprise Gazeta do Povo (2018).

## 7.2 The Counterfactual Case: The Value of Decision Support

The most significant finding lies in the divergence between the model and human decision-making regarding the right-back, **Fagner**. This discrepancy highlights the potential value of the FCS as a bias-mitigating tool.

Throughout the critical window (minutes 40 to 85), Fagner consistently held the **Maximum Priority (100.0)**, ranking as the #1 candidate for substitution in the team, surpassing all players who were actually removed. The fuzzy inference engine identified a ”Critical Risk” scenario composed of:

1. Consistently low technical performance ( $P_{cum}$ );
2. High defensive exposure against the opponent’s key winger (Eden Hazard).

**Outcome:** The player was not substituted. As a consequence—predicted by the high risk calculated by the system—the player continued to struggle in 1v1 duels Gazeta do Povo (2018) and, validating the disciplinary risk logic, received a yellow card at the 90<sup>th</sup> minute. Had the coaching staff acted on the ”Priority 100” alert, a preemptive substitution could have mitigated the defensive liability on the right flank, offering a tactical advantage that human intuition failed to secure in real-time.

### 7.3 External Validation via Expert Analysis

To qualitatively validate the fuzzy rule base, the model’s outputs were compared with post-match ratings provided by specialized sports media. A survey by *Gazeta do Povo* Gazeta do Povo (2018) assigned the lowest ratings in the squad to the exact four players flagged by the FCS: Fagner (4.0), Willian (3.0), Paulinho (3.5), and Gabriel Jesus (2.5).

This convergence between the algorithmic real-time assessment and the *a posteriori* human analysis confirms that the selected fuzzy variables—Fatigue, Momentum, and Positional Risk—serve as robust proxies for player performance and tactical viability.

### 7.4 Cross-Team Validation: The Belgian Context

To assess the model’s generalizability, the Fuzzy Control System was applied to the opponent’s squad. This comparison provides insight into how the algorithm evaluates ”decisive but low-volume” performances compared to ”ineffective” ones.



Figure 4: (each point represents the interval up to the next, e.g., 0 = 0–5 min, 5 = 5–10 min, 10 = 10–15 min)

#### 7.4.1 The Lukaku Paradox: Single Action vs. Prolonged Stagnation

The analysis of Romelu Lukaku offers a critical validation of the system’s ability to prioritize consistent engagement over isolated highlights. Although the striker provided the assist for the second goal at the 31<sup>st</sup> minute, his performance metrics plummeted immediately thereafter. Post-match data reveals that Lukaku recorded only 14 ball touches in 87 minutes of play Revista Época (2018)—a critically low volume that implies long periods of tactical absence.

**Model Diagnosis vs. Human Latency:** The Fuzzy Control System operated as a strict performance auditor, detecting a severe drop in contribution that the human manager tolerated for nearly an entire half.

- **Overpowering the Protection Rule:** Typically, an assist triggers a "Protection Rule" (reducing substitution priority). However, Lukaku’s cumulative performance ( $P_{cum}$ ) was so low, and his participation so scarce, that the negative logic—specifically **Rule R08** (Ineffective Forward) and **Rule R02b** (Critical Fatigue/Decay)—overwhelmed the assist bonus.

- **Early Detection (35' vs 87'):** The model identified Lukaku as the **#1 Substitution Priority** in the squad starting as early as the 35<sup>th</sup> minute (immediately post-assist).
- **Escalation to Maximum Priority:** From that moment on, his priority score grew continuously, reaching the maximum value of **100.0** by the time of 75 minutes, which was the last recorded period in which he had any on-field activity. He was finally substituted at the 87<sup>th</sup> minute.

This timeline demonstrates a significant divergence: while the coach waited until the final minutes to make the change, the algorithm flagged the effective cessation of the player’s contribution over 50 minutes earlier. The substitution at 87’—the first voluntary change by Belgium—validated the model’s diagnosis that, despite the assist, the player had become the most ineffective element on the field due to extreme disconnection from the game flow.

Table 5: Comparative Analysis of Low-Volume Strikers (Brazil vs. Belgium).

Player	Ball Touches	Key Contribution	FCS Priority Rank	Dominant Rules
Gabriel Jesus (BRA)	12	None	#2 (High)	R08 (Inefficacy)
Romelu Lukaku (BEL)	14	1 Assist	#1 (Max)	R08 (Inefficacy)

#### 7.4.2 Nacer Chadli and Physical Limits

Nacer Chadli was substituted at the 83<sup>rd</sup> minute due to injury. This was the only substitution not anticipated by the model, highlighting a key limitation: while the system can effectively monitor fatigue, workload, and signs of physical stress, it cannot predict sudden injuries, which are inherently unexpected and occur independently of observable fatigue indicators. Chadli’s forced exit serves as a reminder that even the most advanced predictive models must account for the unpredictable nature of sports, where accidents and acute injuries can alter the course of a match despite optimal conditioning and monitoring.

## 8 Comparative Analysis: Fuzzy Inference vs. Machine Learning Baselines

To position the proposed Fuzzy Control System (FCS) within the broader landscape of soccer analytics, it is necessary to contrast its performance and utility against purely data-driven Machine Learning (ML) approaches. Recent literature, specifically the comprehensive study by Mohandas et al. (2023), provides a robust baseline for this comparison.

### 8.1 The Predictive Ceiling of Machine Learning

Mohandas et al. (2023) analyzed a massive dataset of 51,738 substitutions across top European leagues, employing supervised learning algorithms including Random Forest, SVM, and Logistic Regression. Their state-of-the-art model (Random Forest) achieved a prediction accuracy of **70%**.

While statistically significant, this result reveals a critical "predictive ceiling." Nearly 30% of substitution decisions remain inexplicable to standard ML models. This gap suggests that substitutions are not purely deterministic events driven by raw data, but are heavily influenced by context, intuition, and tactical nuance—factors that "black-box" models struggle to encode. Furthermore, supervised ML models in this domain suffer from a fundamental conceptual flaw: they are trained to *mimic* the human coach. If the training data contains biased or suboptimal decisions (e.g., a coach delaying a necessary substitution), the ML model learns to replicate that error.

### 8.2 The Explainability Advantage

The primary superiority of the Fuzzy Logic approach lies in its shift from *prediction* to *prescription*. Unlike neural networks or ensemble methods, which output a probability score based on opaque feature interactions, the FCS provides a semantic trace of the decision process.

Table 6 summarizes the structural differences between the two paradigms. While ML

excels at identifying historical patterns, it lacks the ability to justify its output to a human expert. For instance, if an ML model predicts a substitution for a player with high passing accuracy, the coach has no insight into *why*. In contrast, our FCS explicitly outputs the activating rule (e.g., *Rule R04: Rapid Decline in Momentum*), providing an actionable rationale that aligns with coaching language.

Table 6: Structural Comparison: Supervised ML vs. Fuzzy Control System

Dimension	Supervised ML (e.g., Mohandas et al. (2023))	Proposed Fuzzy Control System
Objective	Minimize Prediction Error (Mimic Coach)	Maximize Tactical Utility (Audit Coach)
Handling Uncertainty	Probabilistic (0 to 1 confidence)	Possibilistic (Membership Degrees)
Interpretability	Low (Black Box / Feature Importance)	High (Linguistic Rules)
Response to Outliers	Treats as Noise/Error	Handles via Specific Rules (e.g., R01)
Dependency	Requires massive historical training data	Requires expert knowledge engineering

### 8.3 Case Study: The ”Fagner Paradox” as Proof of Utility

The divergent case of the player Fagner in the Brazil vs. Belgium match (Section 6.2) serves as the definitive empirical proof of the superiority of the Fuzzy approach for *Decision Support*.

- **The ML Prediction:** A supervised model trained on Tite’s historical behavior would likely predict **”No Substitution”** for Fagner, as Tite historically exhibits a conservative bias regarding defensive changes. In an accuracy-based metric (like those in Mohandas et al.), this prediction would be scored as **”Correct”** (True Positive), reinforcing the status quo.
- **The Fuzzy Diagnosis:** Our FCS, unburdened by the need to mimic historical bias, flagged Fagner as **Priority 100.0** based strictly on the logical conjunction of performance metrics and disciplinary risk.

Here, the ”inaccuracy” of the Fuzzy system (disagreeing with the coach) is its greatest asset. By failing to predict the actual event, the system succeeded in identifying the tactical error. While ML models validate the decision that *was* made, the Fuzzy Logic system identifies the decision that *should have been* made. This distinction confirms that for the specific problem of real-time tactical management—where the goal is to

optimize performance rather than predict behavior—the rule-based fuzzy architecture offers a tangible competitive advantage over purely correlational ML baselines.

## 9 Conclusion

This study addresses the critical disconnect between the abundance of granular soccer data and the subjective, high-pressure reality of technical management. By engineering a Fuzzy Logic-based Decision Support System (DSS), we successfully demonstrated that expert heuristics can be codified into a computable framework that operates in real-time. Unlike traditional Machine Learning approaches that aim to *predict* substitutions—ineluctably adopting the biases of historical human decisions—our system provides a *prescriptive* analysis, auditing performance based on objective metrics rather than intuition.

The validation through the 2018 FIFA World Cup case study confirmed the system’s ecological validity and its distinct advantage over purely correlational models. While the system aligned with the coaching staff on obvious performance drops (e.g., Gabriel Jesus), its true value was proven in the divergence. By flagging the right-back position (Fagner) and the opposing striker (Lukaku) as maximum-priority risks significantly earlier than human intervention occurred, or not, the model demonstrated its capacity to mitigate cognitive biases, such as the *sunk cost fallacy* and *outcome bias*, which frequently cloud human judgment in high-stakes environments.

However, limitations remain. The current reliance on event data as a proxy for physical fatigue (`minutesPlayed`) serves as an estimation rather than a direct measurement. Future iterations of this framework will prioritize the integration of biometric telemetry and GPS tracking data to refine the fatigue inputs. Furthermore, we envision the development of a Neuro-Fuzzy (ANFIS) architecture, allowing the rule base to be dynamically calibrated through feedback from professional analysts, effectively creating a hybrid system that learns from expert consensus while maintaining algorithmic objectivity.

Ultimately, this work provides distinct novel contributions to the field of sports analytics: first, by **introducing the first fuzzy-based DSS specifically designed for real-time substitutions**, employing a **hybrid architecture** that contextually modu-

lates disciplinary risk; and second, by **reformulating the PlayeRank metric into a 'Cumulative Mean' metric with role-aware normalization**, effectively eliminating exposure bias to capture intra-match performance decay. By this, proposing a shift from descriptive statistics to actionable, real-time decision support, this system offers coaching staffs a tangible tool to optimize game management when the margin for error is non-existent.

**Reproducibility and Acknowledgments** In the spirit of Open Science, the complete source code, including pre-processing pipelines, fuzzy inference engines, and validation scripts, is publicly available at <https://github.com/Pedro-Passos77/AI-Assisted-Substitution-Decisions-A-Fuzzy-Logic-Approach-to-Real-Time-Game-Management>. We gratefully acknowledge Luca Pappalardo and collaborators for making the *Soccer match event dataset* publicly available (Pappalardo et al., 2019), which was essential for this study. Generative AI tools were employed as drafting assistants during the revision process, under full human supervision, to ensure linguistic accuracy and clarity.

““

## References

CNN Brasil Esportes. Guardiola é o técnico que menos faz substituições na premier league; veja ranking [guardiola is the coach who makes the fewest substitutions in the premier league; see ranking]. <https://www.cnnbrasil.com.br/esportes/futebol/futebol-internacional/guardiola-e-o-tecnico-que-menos-faz-substituicoes-na-premier-league-veja-ranking/>, s.d. Accessed on: 24 Oct. 2025.

ESPN Brasil. Atuações da seleção: Neymar e coutinho se salvam em eliminação, 2018. URL [https://www.espn.com.br/futebol/copa-do-mundo/artigo/\\_/id/4508928/atuacoes-da-selecao-neymar-e-coutinho-se-salvam-em-eliminacao](https://www.espn.com.br/futebol/copa-do-mundo/artigo/_/id/4508928/atuacoes-da-selecao-neymar-e-coutinho-se-salvam-em-eliminacao). Acessado em: Outubro 2025.

ESPN Brasil. Por que guardiola não fez nenhuma substituição no city contra o rb leipzig? ele mesmo responde [why didn't guardiola make any substitutions for city against rb leipzig? he answers]. [https://www.espn.com.br/futebol/manchester-city/artigo/\\_/id/11661194/por-que-guardiola-nao-fez-nenhuma-substituicao-no-city-contra-o-rb-leipzig-ele-mesmo-responde](https://www.espn.com.br/futebol/manchester-city/artigo/_/id/11661194/por-que-guardiola-nao-fez-nenhuma-substituicao-no-city-contra-o-rb-leipzig-ele-mesmo-responde), s.d. Accessed on: 24 Oct. 2025.

ESPN.com.br. Quanto o fluminense ganhou em premiação no mundial de clubes após eliminação para o chelsea na semifinal. [https://www.espn.com.br/futebol/fluminense/artigo/\\_/id/15405687/quanto-fluminense-ganhou-premiação-mundial-de-clubes-após-eliminação-para-o-chelsea-semifinal](https://www.espn.com.br/futebol/fluminense/artigo/_/id/15405687/quanto-fluminense-ganhou-premiação-mundial-de-clubes-após-eliminação-para-o-chelsea-semifinal), s.d. Acesso em: 26 out. 2025.

Gazeta do Povo. Notas do brasil: Fernandinho e gabriel jesus são os piores contra a bélga, 2018. URL <https://www.gazetadopovo.com.br/esportes/copa-2018/notas-brasil-belgica-copa-2018/>. Acessado em: Outubro 2025.

ge. Premiação da premier league: veja quanto cada clube recebeu em 2023/24. <https://ge.globo.com/futebol/futebol-internacional/noticia/2025/02/07/premiação-da-premier-league-veja-quanto-cada-clube-recebeu-em-202324.ghtml>, 2025. Acesso em: 26 out. 2025.

LACCEI. A system for the control of the performance of high level soccer players applying fuzzy logic. In *2023 LACCEI International Multi-Conference for Engineering, Education, and Technology*, 2023. URL [https://laccei.org/LACCEI2023-BuenosAires/papers/Contribution\\_1118\\_a.pdf](https://laccei.org/LACCEI2023-BuenosAires/papers/Contribution_1118_a.pdf).

F. T. Marliere. Sistema de apoio à decisão baseado na lógica fuzzy e aplicado ao futebol de robôs [decision support system based on fuzzy logic and applied to robot soccer]. [https://www2.ufjf.br/eletrica\\_automacao/wp-content/uploads/sites/647/2017/02/TCC\\_Frederick-Tavares-Marliere.pdf](https://www2.ufjf.br/eletrica_automacao/wp-content/uploads/sites/647/2017/02/TCC_Frederick-Tavares-Marliere.pdf), 2017.

Alex Mohandas, Mominul Ahsan, and Julfikar Haider. Tactically maximize game advantage by predicting football substitutions using machine learning. *Big Data and*

*Cognitive Computing*, 7(2):117, 2023. doi: 10.3390/bdcc7020117. URL <https://www.mdpi.com/2504-2289/7/2/117>.

L. Pappalardo. Soccer match event dataset. <https://doi.org/10.6084/m9.figshare.c.4415000.v5>, 2020. [Data set].

L. Pappalardo, P. Cintia, A. Rossi, E. Massucco, P. Ferragina, D. Pedreschi, and F. Giannotti. 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):1–24, 2019. doi: 10.1145/3343172. URL <https://dl.acm.org/doi/10.1145/3343172>.

Revista Época. Gabriel jesus termina a copa sem gols e com menos toques que reservas, 2018. URL <https://epoca.globo.com/esporte/noticia/2018/07/gabriel-jesus-termina-copa-sem-gols.html>. Acessado em: Outubro 2025.

Louise Schmidt, Cristian Lillo, and Javier Bustos. Revisiting playerank, 2024. URL <https://arxiv.org/abs/2410.20038>.

Terra Esportes. Guardiola explica não ter feito substituições no duelo contra o real madrid [guardiola explains not making substitutions in the match against real madrid]. <https://www.terra.com.br/esportes/futebol/internacional/guardiola-explica-nao-ter-feito-substituicoes-no-duelo-contra-o-real-madrid,e3fd37046389eb17727c7e0d1ef1230ahuy6brqn.html>, s.d. Accessed on: 24 Oct. 2025.