Football Player Acquisition & Team Management Optimization

Hongyi Zhan, Oyundari Batbayar, Jiaxuan Wang, Siqi Wang, Yifan Lu

MGSC 662 - DECISION ANALYTICS

McGill University

January 30, 2024

Introduction

In the competitive arena of professional football, strategic player acquisition and management are crucial for a team's success. This project harnesses Gurobi optimization techniques to offer teams a sophisticated, data-driven framework for making informed decisions on player acquisitions, lineups, and coaches. Focusing on multiple objective optimization with a hierarchical approach and discrete optimization through Mixed-Integer Programming (MIP), the project aims to assess player selections and critically evaluate coaching decisions.

1. Optimizing Player Selection and Transfers

Manchester City (2016-2017 Season)

Manchester City, a top-tier English football club, faced a pivotal season in 2016-2017, marked by a transformation under the stewardship of new manager Pep Guardiola. This period represented a shift towards an innovative playing style, and our analysis centered on utilizing optimization techniques to assist Guardiola in this transitional phase. The primary goal was twofold: Firstly, to determine an optimal 25-player squad for registration, aiming to maximize the team's performance while adhering to budgetary constraints, squad composition requirements, and player acquisition limits. Secondly, the focus was on devising the best starting 11 for crucial matches, especially against formidable opponents like Barcelona and Liverpool. By applying optimization models to comprehensive player data, our objective was to strategically strengthen the team through targeted transfers, thereby setting the stage for Guardiola to achieve success in the upcoming seasons.

Manchester United (2016-2017 Season)

In contrast, Manchester United, another Premier League powerhouse, embarked on a different tactical journey under the guidance of coach José Mourinho. Similar to Guardiola, Mourinho's approach involved a blend of tactical formations, notably shifting between 4-2-3-1 and 4-3-3.[1] However, while Guardiola's style is characterized by an emphasis on possession, Mourinho's approach is more defensively oriented and geared towards exploiting counter-attacking opportunities. Our analysis with United focused on optimizing a 23-player selection to cater to Mourinho's diverse tactical needs. This involved a thorough assessment of player attributes to enable flexibility in formation and strategy, essential for adapting to different match scenarios. Such an approach was not only about selecting the best players but also about forming a squad capable of executing Mourinho's vision.

2. Coaching Decision-Making in Team Management

Building on the foundation of these strategically selected teams, we then delve into the significant impact of coaching decisions on team performance. The role of the coach extends beyond selecting the right mix of players; it involves making critical decisions that can unlock the team's full potential. This part of our analysis shifts the focus to the German national football team, exploring how coaching methodologies and tactical decisions can maximize the team's potential. We examined key tournaments, including the 2018 World Cup and Euro 2021, to evaluate the effectiveness of Germany's preferred formations like 3-4-3 and 4-3-3.[2] By constructing multi-objective functions and comparing these formations against actual game decisions, our goal was to uncover the strengths and weaknesses inherent in these tactical choices. This comprehensive analysis aimed not only to shed light on the tactical acumen of the German coaching staff

---

[1] 4-2-3-1 is 4 defenders, 2 defensive midfielders, 3 attacking midfielders, and 1 striker; whereas 4-3-3 is 4 defenders, 3 midfielders, and 3 forwards

[2] Formations are explained in the problem formulation.

but also to identify opportunities for strategic improvements in their approach to international football.

Problem Description and Formulation

Manchester City team

We employed Gurobi to aid Manchester City in forming a team for the 2016-2017 season, focusing on assembling a 25-player squad and crafting match-specific lineups with 11 players against Barcelona and Liverpool. Challenges include balancing team performance within budget constraints and considering player traits and tactical needs against varied playing styles.

The objective function for selecting 25 players for registration

$$\text{Maximize} \quad \sum_{i \in \text{man\_city\_players}} \text{O}[i] \cdot x_i + \sum_{i \in \text{other\_teams\_players}} (\text{O}[i] + \text{A\_s}[i]) \cdot y_i \tag{1}$$

where x_i and y_i are binary decision variables indicating whether player i is selected for the team from Manchester City players and other teams' players, respectively. O[i] and A_s[i] represent the potential rating[3] and adaptation score[4] of player[i], respectively.

Constraints

- Player Acquisition Budget: There is a cap on the amount to be spent on acquisitions. $\sum_{i \in \text{other\_teams\_players}} y_i \cdot \text{V}[i] \leq 150 million$ where y_i is a binary decision variable indicating whether player i from other teams is selected, and V[i]: value_eur$_i$ represents the acquisition cost in euros of player i.

- Team Composition: This constraint ensures that the number of registered players should be 25. $\sum_{i \in \text{man\_city\_players}} x_i + \sum_{i \in \text{other\_teams\_players}} y_i = 25$ along with positional constraints to ensure a well-balanced team with at least 8 defenders, 4 forwards, 4 midfielders, and 2 goalkeepers. $\sum_{i \in \text{man\_city\_defenders}} x_i + \sum_{i \in \text{other\_teams\_defenders}} y_i \geq 8$; $\sum_{i \in \text{man\_city\_forwards}} x_i + \sum_{i \in \text{other\_teams\_forwards}} y_i \geq 4$; $\sum_{i \in \text{man\_city\_midfield}} x_i + \sum_{i \in \text{other\_teams\_midfield}} y_i \geq 4$; and $\sum_{i \in \text{man\_city\_gk}} x_i + \sum_{i \in \text{other\_teams\_gk}} y_i \geq 2$.

- Traded Players: $\sum_{i \in \text{other\_teams\_players}} y_i \leq 4$

- Player Acquisition: The probability of players being bought gets higher as their contract expiration dates get closer. $y_i \leq \text{C}[i]$ where C[i]: club_expiry_soon[i] is a binary parameter that is 1 if the contract of player i is expiring soon (2016,2017, and 2018) and 0 otherwise.

Objective Functions for Liverpool vs Manchester City

Objective	Formula	Priority
Maximize Skill Ball Control	max $\sum_{i \in S} x_i \cdot \text{S\_b}[i]$	2
Maximize Defending Work-rate	max $\sum_{i \in S} x_i \cdot \text{D\_w}[i]$	1
Maximize Power Strength	min $\sum_{i \in S} x_i \cdot \text{Stre}[i]$	0

Table 1: Objective Functions for Liverpool vs Manchester City

---

[3]These variables/scores have already been calculated by sofifa.com. Various studies, including this https://www.kaggle.com/code/kevinmh/evaluating-the-sofifa-com-rating-calculator predicts that their calculation accuracy is good.

[4]Note the reasoning behind this will be explained in detail in the numerical implementation part

where x_i is a binary variable that indicates if player i is selected, and S represents the set of all players under consideration for selection in the team. S_b[i]: skill_ball_control$_i$ is the ball control skill of player i. D_w[i] and Stre[i] defending work rate and power strength of player i.

Team Composition Constraint for 4-2-3-1:

The team selection is subject to several strategic constraints to align with conventional soccer formations.[5] The team is limited to having only one goalkeeper. The defense must be composed of a specific number of players in positions such as Left Back (LB), Right Back (RB), Left Wing-Back (LWB), Right Wing-Back (RWB), Right Center Back (RCB), and Left Center Back (LCB). If a Left Wing-Back is chosen, then a Right Wing-Back and one Center Back are required, but no additional Left or Right Backs can be included, and vice versa. There must be precisely one Central Attacking Midfielder (CAM) in the lineup. The team must have at least one but no more than two Defensive Midfielders (DM). A minimum of two midfielders are needed in total. Moreover, there must be a balance between the players occupying the left and right forward positions, ensuring an even distribution of players across these offensive roles.

Objective Functions for Barcelona vs Manchester City

Objective	Formula	Priority
Maximize Defending Work-rate	$\max \sum_{i \in S} x_i \cdot \text{D\_w}[i]$	2
Maximize Ball Control	$\max \sum_{i \in S} x_i \cdot \text{S\_b}[i]$	1
Maximize Power Stamina	$\min \sum_{i \in S} x_i \cdot \text{Sta}[i]$	0

Table 2: Objective Functions for Liverpool vs Manchester City

Team Composition Constraint for 4-3-3:

There must be only one goalkeeper on the team. The team must have a specific number of defenders, which include left back (LB), right back (RB), left wing-back (LWB), right wing-back (RWB), right center-back (RCB), and left center-back (LCB). The constraints ensure that if the team includes a LWB, then it must also have a RWB and one CB, but it cannot have both an LB and an RB. There must be at least two midfielders in the team. There needs to be a balance between the left and right forward positions. This balance is achieved by ensuring that the total number of players in the left forward positions (left winger LW, left forward LF) is equal to the total number of players in the right forward positions (right winger RW, right forward RF).

Manchester United team

We then explored Manchester United's player composition for the 2016-2017 season. To form a 23-player squad, we needed to select four new players for the team before the season. Our project investigated the historical coaching style of José Mourinho, particularly focusing on his exceptional tenure at Real Madrid. By integrating his historical formation styles and the importance of attributes for various positions, we aimed to maximize the team's performance, all while ensuring that certain constraints are met.

Objective Function	Formula	Priority
Overall Score of The Team	$\sum_{i\in\text{Players}} x_i \cdot \text{overall}_i$	5
Attacking Midfielders/Wingers	$\max \sum_{i\in\text{AMF}\cup\text{Wingers}} x_i \cdot (\sum_j \text{Attacking\_skills}_{i,j})$	4
Strikers	$\max \sum_{i\in\text{ST}} x_i \cdot (\sum_j \text{Strikers\_skills}_{i,j})$	3
Defensive Midfielders	$\max \sum_{i\in\text{ST}} x_i \cdot (\sum_j \text{Defensive\_skills}_{i,j})$	2
Defenders	$\max \sum_{i\in\text{DF}} x_i \cdot (\sum_j \text{Defenders\_skill}_{i,j})$	1
Goal Keepers	$\max \sum_{i\in\text{GoalKeeper}} x_i \cdot (\sum_j \text{GoalKeeper\_skills}_{i,j})$	0

Table 3: Objective Functions for selecting 4 other players to form a 23-player squad with 4-2-3-1 formation

Multi-Objective Functions for Finalizing a 23-Player Sports Squad

Attacking skills encompass dribbling, attacking crossing, skill curve, attacking short passing, mentality vision, and movement agility. Strikers' skills focus on shooting, attacking finishing, pace, mentality composure, attacking heading accuracy, skill ball control, and movement reactions. On the defensive front, skills include defending, mentality interceptions, passing, mentality aggression, physicality, power stamina, and mentality positioning. Defenders are skilled in defending marking awareness, power strength, mentality composure, movement reactions, defending standing tackle, and defending sliding tackle. Lastly, a goalkeeper's skill set is defined by goalkeeping diving, goalkeeping handling, goalkeeping positioning, goalkeeping reflexes, and goalkeeping kicking[6].

Constraints For The Main Model

- Transfer Budget: The total market value of selected players does not exceed 200 million euros, maintaining financial discipline in team selection. $\sum_{i\in\text{Players}} x_i \cdot \text{value\_eur}_i \leq \text{Budget}$

- Wage Budget: The total weekly wages of selected players to the club's wage budget, ensuring financial sustainability based on the club's revenue. $\sum_{i\in\text{Players}} x_i \cdot \text{wage}_i \leq \text{Budget}$

- Non-EU Player: Limits the number of non-EU players to a maximum of 17, ensuring compliance with league or tournament regulations. $\sum_{i\in\text{NonEUPlayers}} x_i \leq \text{MaxNonEU}$

- Goalkeeper: Exactly one goalkeeper is selected for the team, adhering to the standard football team composition. $\sum_{i\in\text{Goalkeepers}} x_i = 1$,Goalkeepers $= \{gk \mid gk \in \text{Players}, \text{'GK'} \in \text{player\_positions}\}$

- Select 4 New Players: Exactly 4 new players are selected, allowing for strategic team refreshment or strengthening. $\sum_{i\in\text{Players}} x_i = 4$

- Age: The average age of selected players at or below 27 years, helping to maintain a balanced team in terms of age. $\sum_{i\in\text{Players}} x_i \cdot \text{age}_i \leq \text{MaxAge} \cdot \text{TotalPlayers}$

Constraints For Further Analysis

After selecting a 23-player squad, we attempted to choose different starting lineup combinations from within these 23 players. Taking into account Mourinho's preferred styles, we selected one group in a 4-2-3-1 formation and another in a 4-3-3 formation.[7] The objective function in

---

[5]Similar mathematical formulations are used in the German Coaching Decision Part, so skipping the mathematical notation here and all should be included in the code

[6]See Appendix Table 16: Player Abilities and Attributes

[7]See Appendix Table 17: Skills Required for Different Positions in a 4-2-3-1 Formation; And Table 18: Skills Required for Different Positions in a 4-3-3 Formation

this process is similar to the main model, focusing on maximizing the skills of players in each position, which we will not elaborate on further. Here, we introduced the following additional constraints to get the 11 starting line-up.

- Each player is selected for only one position. $\sum_{\text{pos}\in\text{Positions}} \text{select}[p, \text{pos}] \leq 1 \quad \forall p \in \text{Players}$

- Position Constraint: Positions = $\{pos \mid pos \in \text{Players}, \text{Pos\_Abbreviation} \in \text{player\_positions}\}$
  $\sum_{i\in\text{Positions}} x_i = \text{NumRequiredForPosition}$

Positions	Number Required For Position
GoalKeepers	GoalKeepers (GK = 1)
Defenders	Center Backs (CB = 2), Left Backs (LB = 1), Right Backs (RB = 1)
Defensive Midfielder	Defensive Midfielder (CDM = 1)
Attacking Midfielder	Central Attacking Midfielder (CAM = 1), Left/Right Midfielder (LM/RM = 1), Left/Right Winger (LW/RW = 1)
Striker	Striker (ST = 1)

Table 4: Number of Players Required for Position

Coaching Decision-Making in Team Management

Now, to thoroughly evaluate the effectiveness of a coach's decision-making, we initiated our study by tackling the complex challenge of optimizing the squad selection and formation strategy for the German national football team during the 2018 World Cup. Initially, a 23-man squad was selected from among German players based on their performance across 27 positions, ensuring a comprehensive evaluation of versatility and skill. Subsequently, the focus shifted to setting stricter positional limits within the squad to facilitate the formation of different starting 11 lineups.

Objective Function for choosing 23 players in 2018

$$\text{Maximize} \quad \frac{1}{23}\sum_{i=0}^{\text{num}-1}\sum_{j=0}^{26} n_{ij}\cdot\text{ability}_{ij} \tag{2}$$

where ability$_{ij}$ shows the ability for plays i in position j. $n_{ij} \in \{0, 1\}$ and num is the total number of German players candidate. The index for each position j can be seen in Table 19.

1: Player Assignment Constraints

These constraints ensure that each player is assigned to at most one position and that the total number of players selected is exactly 23. We can write the form like: $\forall i \in \{0,\ldots,\text{num}-1\}$, $\sum_{j=0}^{26} n_{ij} \leq 1$ and $\sum_{i=0}^{\text{num}-1}\sum_{j=0}^{26} n_{ij} = 23$.

2: Position-Specific Constraints

These constraints ensure a balanced team composition by mandating exactly 3 goalkeepers for a strong defense $\sum_{i=0}^{\text{num}-1} n_{i,26} = 3$, at least 2 forwards for offensive power $\sum_{i=0}^{\text{num}-1} n_{i,1} + \sum_{i=0}^{\text{num}-1} n_{i,5} \geq 2$, and a minimum of one player each for left back (LB) and right back (RB) to secure both flanks $\sum_{i=0}^{\text{num}-1} n_{i,k} \geq 1$, k $\in \{21, 25\}$. Crucially, the model also includes foot preference constraints for (LB/LWB) and (RB/RWB), ensuring that players in these positions are optimally suited for their roles. $\forall i \in \{0,\ldots,\text{num}-1\}$, $n_{i,k} \leq \text{foot}_i$, k $\in \{16, 20, 21, 25\}$, $\text{foot}_i \in \{0, 1\}$ Additionally, the model requires at least 2 right center backs (RCB) and 2 left center backs (LCB) for a solid central defense $\sum_{i=0}^{\text{num}-1} n_{i,k} \geq 2$, k $\in \{22, 24\}$.

3: Team Composition Constraints

These constraints ensure a well-balanced team, with a backfield of 7 to 9 players for defense, and a frontfield of 3 to 5 players for offense. They also maintain a strategic balance between left and right-sided players, with a maximum difference of two players. $|l - r| \leq 2$ where $l = \sum_{i=0}^{\text{num}-1} \sum_j n_{ij}$, $j \in \{0, 3, 4, 8, 11, 12, 16, 17, 21\}$ and $r = \sum_{i=0}^{\text{num}-1} \sum_j n_{ij}$, $j \in \{2, 6, 7, 10, 14, 15, 19, 20, 25\}$. Additionally, no position can have more than 3 players $\forall j \in \{0, \ldots, 26\}$, $\sum_{i=0}^{\text{num}-1} n_{ij} \leq 3$, ensuring diversity and tactical versatility across the team.

Choosing staring 11 lineups

To adapt for an 11-player soccer team, we changed the objective function above to optimize the performance for the 11 line-up. Key changes for constraints include limiting the team to one goalkeeper, specifying exact numbers for defenders like (LB, RB, LWB, RWB, RCB, and LCB), to guarantee if the team has LWB then must have RWB and one CB but no more LB and RB vice versa. $\sum_{i=0}^{\text{num}-1} n_{i,16} = \sum_{i=0}^{\text{num}-1} n_{i,20}$ and $\sum_{i=0}^{\text{num}-1} n_{i,16} = \sum_{i=0}^{\text{num}-1} n_{i,23}$. Additionally, it mandates a minimum of two midfielders $\sum_{i=0}^{\text{num}-1} (n_{i,12} + n_{i,13} + n_{i,14} + n_{i,17} + n_{i,18} + n_{i,19}) \geq 2$ and balance between left and right forward positions rf = lf where rf $= \sum_{i=0}^{\text{num}-1} (n_{i,6} + n_{i,7} + n_{i,15})$ lf $= \sum_{i=0}^{\text{num}-1} (n_{i,3} + n_{i,4} + n_{i,11})$, aligning with standard soccer team structures.

Multiple Objective Functions for choosing 26 players & 11 lineups in 2021

The key issue of Germany's unexpected exit from the 2018 World Cup was identified as low attacking efficiency. Considering this, a strategic approach to augment the team's offensive dynamics was adopted for the Euro 21 squad selection and the starting eleven. This was achieved by modifying the objective functions in the selection process, specifically to enhance attacking attributes, while keeping the constraints consistent with the previous model. Additionally, a detailed comparison was made between the conventional 4-3-3 formation and the coach's 3-4-3 formation, aimed at identifying the most effective strategy to strengthen Germany's offensive capabilities and overall team performance for the upcoming tournament.

Objective Function	Formula	Priority
Attacking	$\sum_{i=0}^{\text{num}-1} \text{an}_i \cdot \text{attack}_i$	3
Defending Heading	$\sum_{i=0}^{\text{num}-1} \text{dh}_i \cdot \text{defend\_heading\_list}_i$	2
Potentialing	$\frac{1}{26} \sum_{i=0}^{\text{num}-1} \text{potn}_i \cdot \text{potential}_i$	1
Total	$\frac{1}{26} \sum_{i=0}^{\text{num}-1} \sum_{j=0}^{26} n_{ij} \cdot \text{ability}_{ij}$	0

Table 5: Objective Functions for selecting 26 players in 2021

Objective Function	Formula	Priority
Defensive Awareness	$\sum_{i=0}^{\text{num}-1} \text{den}_i \cdot \text{new\_defawareness}_i$	0
Tackling	$\sum_{i=0}^{\text{num}-1} \text{tan}_i \cdot \text{new\_tackle}_i$	1
Positioning	$\sum_{i=0}^{\text{num}-1} \text{pon}_i \cdot \text{new\_positioning}_i$	2
Control	$\sum_{i=0}^{\text{num}-1} \text{con}_i \cdot \text{new\_control}_i$	3
Stamina	$\sum_{i=0}^{\text{num}-1} \text{stn}_i \cdot \text{new\_stamina}_i$	4
Crossing	$\sum_{i=0}^{\text{num}-1} \text{cn}_i \cdot \text{new\_cross}_i$	5
Total	$\frac{1}{11} \sum_{i=0}^{\text{num}-1} \sum_{j=0}^{26} n_{ij} \cdot \text{new\_26\_ability}_{ij}$	6

Table 6: Objective Functions for selecting starting 11 lineup with 3-4-3 formation

where attack$_i$, Defending Heading$_i$ and Defending Heading$_i$ shows the ability of a player to score, the accuracy when using the head to pass or clear the ball, and player's destined growth

peak for players i. Defensive Awareness, Tackling, Positioning, Control, Stamina and Crossing shows the ability to track and defend an opposing player, perform standing tackle, spot open space and move into good positions, keep possession of the ball, sustain prolonged physical and the quality and accuracy of a player's crosses. num is the total number of German players. $an_{ij}, dh_i, potn_i, den_i, tan_i, pon_i, con_i, stn_i, cn_i, n_{ij} \in \{0, 1\}$.

Numerical Implementation and Results

We used FIFA data from 2014-2023.[8] Pre-processing steps are described in the Appendix.

Manchester City team

In choosing the players, we examined key constraints like the 25-player squad limit and UEFA's financial fair play guidelines. Our approach in selecting the 25 was simple yet realistic. Like all reputable clubs, City would need at least 8 defenders (4 CB, 2 LB, 2RB), 4 midfielders (combination of CAM/CM/CDM), 4 forwards (1 ST, 1 CF, 1 RW, 1 LW), and 2 goalkeepers.

Our analysis focused on Pep Guardiola's 2016/17 season in charge after his renowned success at Barcelona. Knowing Guardiola's preference for pass-and-control tactics, we looked at which players already understood his system or had experience in similar styles. When choosing players to acquire[9], the "adaptation score" in our first objective function gave an advantage to candidates from Guardiola's former clubs or teams like Barcelona that emphasize possession and pressing from the front.[10] The 2016 summer transfer window was our scenario when Premier League sides transfer players. Using FIFA player data from before the summer of 2016, we identified potential targets with contracts expiring in 2016, 2017, and 2018.[11]

Looking ahead to key 2016/17 matches, we profiled opponents like Barcelona with possessional style and short passes and Liverpool with a more direct approach bypassing midfield. Through historical data analysis, we concluded City would be better off against Liverpool having good defense, ball control, and players with strength, and against Barcelona having good defense, ball control, and high stamina. Guardiola's famous 4-3-3 and 4-2-3-1 formations were also included.

Results: Were there any significant differences between the actual selection and Gurobi's results?

When playing against Liverpool:[12]
Nicolás Otamendi, Pablo Zabaleta, Yaya Touré, Kevin De Bruyne, David Silva, and Sergio Agüero were key figures featured in both the actual Manchester City lineup and the modeled selection, representing the team's spine with a blend of defensive solidity and mastery during the clash against Liverpool. The model's selection for Manchester City contrasts with the actual team by offering a fusion of seasoned expertise and exceptional talent that could have enhanced their chances against Liverpool. Joe Hart in goal and Vincent Kompany in defense would have brought invaluable experience and leadership, possibly outshining Claudio Bravo's then-recent acquisition and John Stones' developing role. The choice of Neymar adds a world-class attacking dimension absent from the actual lineup, potentially surpassing the impact of the actual

---

[8]See Appendix for Data Description

[9]Note that due to data restriction, the study is limited to only buying scenarios

[10]FC Bayern, FC Barcelona, Borussia Dortmund, Liverpool, Tottenham Hotspur, RB Leipzig, RB Salzburg, Ajax, Bayer Levrkusen, CD Lugo, Las Palmas, Genoa, Atalanta, Athletic Bilbao, Vitesse, Borussia Monchengladbach, Nice. The calculation is described in Code file.

[11]According to https://www.transfermarkt.us/statistik/endendevertraege, players with expiring contracts are more likely to be bought/acquired by other teams.

[12]Result Tables for both Barcelona and Liverpool's matches are in the Appendix

wingers with his unpredictable play. This hypothetical lineup suggests a side equipped with a robust defensive core, complemented by a more varied and possibly more explosive offensive capability compared to the actual selection. In the actual match, Manchester lost 0-1.

When playing against Barcelona:
The addition of Neymar to the right wing and the astute placement of S. Agüero as the striker in our well-constructed Manchester City squad may have been game-changers. David Silva's skill in centre attacking, along with K. De Bruyne in left midfield, would have improved the team's tactical approach. A strong middle defence might be supplied by N. Otamendi and V. Kompany. The expertise of J. Hart in goal and the presence of Y. Touré in centre midfield would strengthen the team's spine. The team's structure is set up to take use of both offensive creativity and defensive stability, with P. Zabaleta and G. Clichy flanking each other in their full-back positions. In the actual match, Manchester lost 0-4 to Barcelona.

Manchester United team

In choosing the 23 players for Manchester United, our approach was a little different than Manchester City. We first reviewed the roster[13] as it stood before the start of the 2016/17 season and focused on identifying additional 4 players from our extensive dataset. The dynamic scenario involved in this problem necessitated the need for constructing multi-objective functions. Our first objective function was to maximize the sum of the overall scores of all players on the team subject to different requirements for formations and positions.

To understand the formations and positions better, we explored coach Mourinho's exceptional performance at Real Madrid and found that he frequently used the 4-2-3-1 formation which offered a balanced mix of defensive and attacking talents. With this in mind, to construct the multi-objective functions, we considered key skills needed for each position in the 4-2-3-1 formation.[14] For instance, for defensive midfielders, we focused on indispensable skills such as defending, passing, physicality, and positioning during the transition between defense and attack, summing up the related skill scores of each player to obtain the 4th objective function. For the striker position, we used variables like shooting techniques, speed, and reaction times which allow players to maintain dominance in confrontations, and summed them in the 3rd objective function. Similarly, we carefully designed corresponding objective functions for different positions, assigning different priorities based on their importance in Mourinho's tactical system. The priority order of our objective functions was as follows: Overall Score of The Team, Attacking Midfielders/Wingers, Strikers, Defensive Midfielders, Defenders, and Goalkeepers. This order reflects our strategic emphasis and alignment with Mourinho's tactical preferences. This approach not only allowed us to optimize the overall performance of the team but also ensured that the team's configuration reflected Mourinho's tactical arrangement in the 4-2-3-1 formation. Through the precise settings of these objective functions, our model effectively simulated the ideal combination of players to meet the on-field demands of specific tactics.

In setting the constraints for this model, we focused on ensuring that the selected players adapted to the 4-2-3-1 formation. These constraints included transfer budget limits (Constraint 1)[15], wage limits (Constraint 2)[16], specific requirements for the number of players in particular

---

[13]Identified 19 active players, https://m.football-lineups.com/team/Manchester-United/Europa-League-2016–2017/players

[14]See Appendix Table 16: Player Abilities and Attributes

[15]https://www.express.co.uk/sport/football/694314/Manchester-United-latest-transfer-news-Jose-Mourinho-Premier-League

[16]https://www.manchestereveningnews.co.uk/sport/football/football-news/man-utd-wages-players-salaries-17166985

positions (Constraint 4), age requirement (Constraint 6)[17], and so on. By implementing these constraints in Gurobi, we were able to explore the optimal combination of players while maintaining a sense of realism.

Results: Who were the new players selected with Gurobi?

Gurobi's optimization model has chosen an impressive lineup of players for the team, each with distinct capabilities and roles.[18] Cristiano Ronaldo, a forward known for his incredible skill, speed, and goal-scoring ability, brings a high level of offense to the team. David De Gea, a goalkeeper renowned for his remarkable reflexes and shot-stopping abilities, significantly strengthens the team's defense. Giorgio Chiellini, a central defender, is known for his aggressive tackling and leadership in the backline, enhancing the team's defensive solidity. Lastly, Kylian Mbappé, a striker with exceptional pace and dribbling skills, adds to the team's attacking threat. This selection presents a balanced mix of defense and attack, improving the team's overall performance on the field.

What were the objective function values?

The optimization outcomes reveal a strategic focus across various objectives. Objective 0 (Maximize Overall Score) scored 356.0, indicating that the average overall score of the 4 new players is 89, which is a quite satisfying result. Objective 2 (Maximize Striker Skills) shows the highest value of 1228.0, underscoring the model's emphasis on enhancing the team's offensive strength. Objective 1 (Maximize Attacking Midfielder Skills) reached a significant value of 513.0, reflecting a concerted effort to boost midfield creativity. Objective 4 (Maximize Defender Skills), with a value of 529.0, highlights the importance placed on defensive robustness. Objective 5 (Maximize Goalkeeper Skills) achieved a notable value of 438.0, indicating a key investment in the goalkeeper's role. In contrast, Objective 3 (Maximize Defensive Midfielder Skills) scored 0.0, showing a balanced approach to overall team skill and a lesser focus on defensive midfield capabilities. The model's use of a 239.5 million euro budget well within our 200 million pounds (about 240 million euro) limit, showcasing prudent financial management in aligning with the strategic goals. **Comparing our model selection** with Manchester United's actual signings for the 2016-2017 season reveals both similarities and distinct focuses. Both choices emphasize a balance between attack and defense. Our selection of forwards like Cristiano Ronaldo and Kylian Mbappé, known for their offensive prowess, mirrors United's acquisition of Zlatan Ibrahimović. However, there's a divergence in midfield strategy; unlike United's signing of versatile midfielder Paul Pogba, our model didn't prioritize a similar style player. This difference likely stems from our model's focus on optimizing specific tactical styles and statistical attributes, contrasting with the club's real-world considerations. This comparison offers insights into the varying approaches of data-driven models versus traditional club management in team composition.

The Starting Line-Up of Different Formations (4-2-3-1 And 4-3-3)

After we explored the adaptability of a 23-player Manchester United squad, we initially optimized for Jose Mourinho's preferred 4-2-3-1 formation, and then reconfigured for a 4-3-3 setup[19]. The transition from a balanced 4-2-3-1, with a strong defensive backbone and a dynamic attack led by Zlatan Ibrahimović and Cristiano Ronaldo, to an attack-oriented 4-3-3, added Kylian

---

[17]https://www.encyclopedia.com/sports/sports-fitness-recreation-and-leisure-magazines/aging-and-athletic-performance

[18]See Appendix Table 14: 23-Player Squad of Manchester United

[19]See Appendix Table 15: Starting Line-Up Squad Selection (For 4-2-3-1 and 4-3-3 Formation)

Mbappé as an extra striker, showcasing the versatility and strategic depth of the squad. This analysis not only reflects the tactical flexibility of the team but also underscores the importance of player adaptability in modern football, offering valuable insights for strategic planning under varying tactical demands.

Coaching Decision-Making in Team Management

2018 World Cup & Euro 2021 Cup Analysis

Here, we crafted our objective function to maximize the overall team value, which hinges on the aggregated performance metrics of 23 players across 27 positions, as referenced in Equation 2. Our model incorporated meticulously designed constraints to emulate real-world team compositions, addressing the distribution of player positions and the equilibrium of footedness.

The results, detailed in Table 7, facilitated a comparative analysis with the actual team configurations. This comparison was crucial in assessing the alignment of our optimized strategy with the coach's real-world decisions for the top 23 players, thereby validating the practicality of our approach in team selection. Extending our analysis, we used the chosen 23 players to formulate optimal starting 11 lineups. Initially adopting a formation-agnostic approach, we instructed Gurobi to optimize the team structure, leading to a 4-3-3 formation (Fig.2). A subsequent model explored a 3-4-3 formation (Fig.3), featuring more stringent constraints, especially for positions like LWB and RWB. This methodology enabled a thorough evaluation of diverse strategic formations against the real German lineup in the 2018 World Cup.

To address Germany's tactical shortcomings, particularly in offensive efficiency, [20] we introduced multiple objective functions (Table.5) focusing on enhancing attacking capabilities for choosing the 26-player squad. Priorities include maximizing scoring proficiency, aerial accuracy for defensive players, and identifying young talents with high growth potential for the 2021 Euro Cup. For the starting 11, we proposed a more aggressive 4-3-3 formation (Fig.3), optimizing speed for wingers and the crossing quality of full-backs. Additionally, we provided insights into a 3-4-3 formation, emphasizing the pivotal roles of LWB and RWB, midfield defensive responsibilities, and attacking creativity. Key attributes like control, spatial awareness, and defensive acumen in central defenders are underscored. Utilizing multiple objective functions (Table.6), we aimed to amplify performance in the 3-4-3 formation and scrutinize the coach's strategic utilization, thereby unlocking the team's full potential.

Results

Squad	Gurobi	Actual
2018 World Cup 23 players	84.91	84.45
2018 World Cup 11 lineup (No priori info)	86.09	84.09
2018 World Cup 11 lineup (3 Back formation)	86	NA
2021 Euro Cup 26 players	84.27	83.12
2021 Euro Cup 11 lineup (4-3-3)	86.36	NA
2021 Euro Cup 11 lineup (3-4-3)	86.27	85.81

Table 7: Objective Value Comparison

---

[20]https://www.bundesliga.com/en/news/Bundesliga/how-germany-could-look-at-the-2018-world-cup-in-russia-447224.jsp

How realistic were Gurobi's predictions?

For the 2018 World Cup, Gurobi opted for different goalkeeping and defensive choices, with an emphasis on speed and adaptability in the forward line. These changes, while notable, only marginally increased the team's overall rating, indicating a slight improvement in team synergy and player ratings.

The Gurobi-optimized formations for the 2018 World Cup, namely 4-3-3 and 3-4-3, show strategic tactical adjustments, aiming to bolster defense and introduce innovation in the attack. These formations, however, only slightly surpassed the actual team's formation in rating, suggesting that while the Gurobi model identified potential improvements, the actual coach's decisions were very close in terms of optimizing team performance.

The comparison with the Euro 2021 Cup squad revealed an even closer alignment between the coach's decisions and the Gurobi model. The Gurobi-optimized formations (4-3-3 & 3-4-3) as shown Fig.7 for this tournament focused on enhancing the forward line while maintaining the core lineup and strategic approach of the actual team. This subtle divergence in starting eleven highlighted the model's focus on fine-tuning rather than overhauling the team strategy.

Overall, the analysis showed that while the Gurobi model proposed certain changes that could potentially enhance team performance, the actual coach's decisions were very much in line with these recommendations, especially for the core lineup. The coach's approach reflected a balanced strategy that weighed player familiarity, real-match dynamics, and statistical optimization. **This indicated that the coach was effectively** optimizing the team's composition and strategy within the dynamic context of top-level football.

Problem Extensions

Manchester City and Manchester United

Utilizing Gurobi, we executed strategic trades to acquire players optimally aligned with City's and United's goals for the 2016-2017 season. Looking ahead, there remains a promising avenue for further refinement. For complex strategic planning, the model can be expanded to a multi-period formulation[21] to optimize lineups and rotations over an entire season or multiple seasons while accounting for uncertainties like player injuries, fatigue levels, and form variations over time. Since league games require steady play over an extended length of time, squad depth and rotation are essential to avoiding player burnout. The model can incorporate player preservation tactics, making certain that star players get breaks when needed and that up-and-coming players are given chances, both of which help to keep the team young and competitive. This would make the model more realistic and useful for long-term strategic planning compared to the current single-tournament focus.

Our Gurobi-based methodology has the ability to advance into real-time analytics, providing a substantial improvement in strategic depth. We may modify our strategy in-game according to player performance, weather, or even the opposition's changing tactics by incorporating real-time data feeds. This degree of dynamism would provide a tactical advantage that might be the difference between a win and a defeat by enabling a reactive approach to replacement patterns

---

[21]Source: Hall, H., Tao, S., Bouncken, R. B., and Puterman, M. L. (2018). A multi-period stochastic programming approach to Major League Baseball roster management. European Journal of Operational Research, 299(3), 1051-1064.

and in-game modifications.

Additionally, the program might forecast future player trajectories by using machine learning techniques, spotting possible breakout stars and declining performances before they show up on the pitch. These forecasts could direct the club's actions in the transfer market, which ensures that its advantage over competitors is maintained. Moreover, including a wider range of performance metrics, such as psychological resilience and team cohesion, may also lead to a more thorough assessment of the squad's capabilities. This would allow for a more sophisticated approach to team development that goes beyond traditional statistical methods.

Coaching Decision-Making in Team Management

The integration of scenario-based strategy modeling and training program optimization into the Gurobi data-driven framework represents a significant advancement in evaluating and enhancing national team coaching performance. Scenario-based strategy modeling is particularly crucial, as it allows coaches to develop tailored tactics for specific match situations such as leading, trailing, or drawing. This component of the model leverages advanced predictive analytics to simulate a variety of game scenarios, equipping coaches with strategic insights necessary for effective in-game decision-making. The ability to adapt to the dynamic nature of matches is fundamental in competitive play, and this strategic planning ensures teams are well-prepared for a range of potential game states. In addition, training program optimization is a key aspect of this approach. This process involves analyzing data from both in-game performance and training sessions to create targeted training programs. These programs focus on enhancing specific skills and strategies that align with the team's identified needs, particularly in frequently encountered match scenarios. This level of customization in training ensures not just physical fitness but also strategic readiness, enabling teams to execute game plans with greater precision.

Recommendations and Conclusions

In conclusion, our sophisticated modeling approaches uniquely cater to the distinct tactical needs of Manchester City, Manchester United, and the German national team. For Manchester City, under Coach Guardiola's guidance, our model advocated a strategic mix of experience and creativity, specifically tailored to counter the challenges of the 2016 season. This approach, crucial against teams like Liverpool and Barcelona, aimed at exploiting their weaknesses and could have been a game-changer in a season plagued by inconsistency. Similarly, for Manchester United, our analysis led to a tactical setup that emphasized adaptability and versatility of attributes, reflecting the club's style and managerial preferences. This strategy, with its foundation in flexible formulation and a focus on sustainable development through young player integration, prepares the team to efficiently adapt to various formations and opponents.

Lastly, for national teams, our enhanced Gurobi model emphasizes strategic flexibility and data-centric decision-making. It adapts strategies to suit international play styles. This approach ensures that national team selections are statistically sound and contextually relevant, offering a robust metric for evaluating coaching effectiveness and aligning with the evolving nature of international football. These models collectively highlight the importance of data-driven strategies in optimizing performance across various football domains.

To extend and enhance these models, we recommend including multi-period planning for better lineup and rotation strategies over seasons and considering player conditions and market dynamics. Real-time analytics could further revolutionize strategy, allowing in-game adjustments based on live data, and machine learning could predict player potential for smarter transfers.

Moreover, integrating scenario-based modeling and training optimizations could significantly improve coaching decisions, tailoring tactics to specific game situations, and optimizing training to meet strategic needs.

Appendix

Data Description

The dataset from kaggle has been scraped from sofifa.com. It contains information on every player, coach, and team available in FIFA 15, 16, 17, 18, 19, 20, 21, 22, and 23. In particular, it contains all FIFA updates from 10th September 2015 until 13th January 2023. The dataset has 32861 rows, 110 attributes for players, 8 attributes for coaches, and 54 attributes for teams. URL of the scraped players, coaches, and teams, along with player positions, and roles in their respective clubs and national teams have been included. Player attributes with statistics such as Attacking, Skills, Defense, Mentality, have been included. Note these values have already been calculated by sofifa.com. Player personal data like Nationality, Club, DateOfBirth, Wage, Salary are also available. Sample exploratory data analysis has been included within the interactive code documentation. Overall, there are no major missing values; most variables related to skill were normally distributed; and skill levels are calculated out of 100.

Manchester City result

Player Name	Position
David Josué Jiménez Silva	CAM, LM
Sergio Leonel Agüero del Castillo	ST
Kevin De Bruyne	CAM, LM, RM
Gnégnéri Yaya Touré	CM, CDM
Vincent Jean Mpoy Kompany	CB
Charles Joseph John Hart	GK
Nicolás Hernán Gonzalo Otamendi	CB
Samir Nasri	LM, RM
Pablo Javier Zabaleta Girod	RB
Raheem Sterling	LM, ST, RM
Martín Gastón Demichelis	CB
Wilfried Guemiand Bony	ST
Gaël Dimitri Clichy	LB
Eliaquim Mangala	CB
Bacary Sagna	RB
Fabian Delph	CM, CDM
Bruno Zuculini	CM, RM
Kelechi Promise Iheanacho	ST
Thierry Ambrose	ST, RW
Angus Gunn	GK
David Faupala	ST
Neymar da Silva Santos Júnior	LW
Mario Götze	CAM, LM, CF, CM
Christian Dannemann Eriksen	CAM, LM
Raphaël Adelino José Guerreiro	LM, LB

Table 8: Player Positions for the selected 25 players

Player Name	Position
Sergio Aguero (Aguero Del Castillo)	CF
Claudio Bravo (Bravo Munoz)	GK
Wilfredo Caballero (Caballero Lazcano)	GK
Gael Clichy (Clichy)	LB
Kevin De Bruyne (De Bruyne)	CM, CAM
Fabian Delph (Delph)	CDM, CM, LB
Manuel Agudo Duran (Nolito)	LW
Ilkay Gundogan (Gundogan)	MF
David Josue (Jimenez Silva)	CAM, CM
Aleksandar Kolarov (Kolarov)	LB, CB
Vincent Kompany (Kompany)	CB
Fernando Luiz Roza (Luiz Roza)	CDM, CB
Jesus Navas (Navas)	RB, RM
Nicolas Otamendi (Otamendi)	CB
Fernando Francisco (Reges)	CDM, CB
Bacary Sagna (Sagna)	RB, RWB
Raheem Sterling (Sterling)	LW, RW
John Stones (Stones)	CB
Yaya Toure (Toure)	CM, CDM
Pablo Zabaleta (Zabaleta Girod)	RB, CDM

Table 9: Player Positions for the Players in Real Life

Player Name	Position
P Zabaleta	RB, CDM
N Otamendi	CB
A Kolarov	LB, CB
Fernandinho	CDM, CB
R Sterling	LW, RW
D Silva	CAM, CM
K De Bruyne	CM, CAM
C Bravo	GK
J Stones	CB
I Gündogan	CM
Nolito	LW

Table 10: Real life Player Position for Game Against Barcelona

Player Name	Position
Neymar	RW
S. Agüero	ST
K. De Bruyne	LM
David Silva	CAM
V. Kompany	RCB
N. Otamendi	LCB
Y. Touré	RCM
J. Hart	GK
C. Eriksen	LCM
P. Zabaleta	RB
G. Clichy	LB

Table 11: Selected life Player Position for Game Against Barcelona

Player Name	Position
C. Bravo	GK
P. Zabaleta	RB, CDM
J. Stones	CB
N. Otamendi	CB
A. Kolarov	LB, CB
Y. Touré	CM, DM, AM
Fernandinho	CDM, CB
R. Sterling	LW, RW
K. De Bruyne	CM, CAM
David Silva	CAM, CM
S. Agüero	CF, SS

Table 12: Real life Player Position for Game Against Liverpool

Player Name	Position
Neymar	LW
S. Agüero	ST
K. De Bruyne	RM
David Silva	CAM
V. Kompany	LCB
N. Otamendi	RCB
Y. Touré	RCM
J. Hart	GK
P. Zabaleta	LDM
B. Sagna	RB
G. Clichy	LB

Table 13: Selected life Player Position for Game Against Liverpool

Manchester United Result

No	Name	Type	Position
1	David de Gea	Original	Goalkeeper (GK)
2	Sergio Romero	Original	Goalkeeper (GK)
3	Eric Bailly	Original	Defenders(CB, LB, RB)
4	Phil Jones	Original	Defenders(CB, LB, RB)
5	Marcos Rojo	Original	Defenders(CB, LB, RB)
6	Chris Smalling	Original	Defenders(CB, LB, RB)
7	Daley Blind	Original	Defenders(CB, LB, RB)
8	Michael Carrick	Original	Defensive Midfiender (CDM)
9	Ander Herrera	Original	Defensive Midfiender (CDM)
10	Paul Pogba	Original	Attacking Midfiender (CAM, LM/RM, LW/RW)
11	Juan Mata	Original	Attacking Midfiender (CAM, LM/RM, LW/RW)
12	Jesse Lingard	Original	Attacking Midfiender (CAM, LM/RM, LW/RW)
13	Ashley Young	Original	Attacking Midfiender (CAM, LM/RM, LW/RW)
14	Adnan Januzaj	Original	Attacking Midfiender (CAM, LM/RM, LW/RW)
15	Zlatan Ibrahimović	Original	Striker (ST)
16	Wayne Rooney	Original	Striker (ST)
17	Anthony Martial	Original	Striker (ST)
18	Marcus Rashford	Original	Striker (ST)
19	Memphis Depay	Original	Striker (ST)
20	Cristiano Ronaldo dos Santos Aveiro	Gurobi	Attacking Midfiender (CAM, LM/RM, LW/RW) , Striker (ST)
21	David De Gea Quintana	Gurobi	Goalkeeper (GK)
22	Giorgio Chiellini	Gurobi	Defenders(CB, LB, RB)
23	Kylian Mbappé Lottin	Gurobi	Striker (ST)

Table 14: 23-Player Squad of Manchester United

Position	4-2-3-1 Formation	4-3-3 Formation
CB	Chris Smalling	Chris Smalling
LB	Daley Blind	Faustino Marcos Alberto Rojo
CDM	Ander Herrera Agüera	Ander Herrera Agüera
CDM	Paul Pogba	-
RM	Juan Manuel Mata García	Juan Manuel Mata García
RB	Ashley Young	Ashley Young
ST	Zlatan Ibrahimović	Zlatan Ibrahimović
CAM	Wayne Mark Rooney	Paul Pogba
LW	Cristiano Ronaldo dos Santos Aveiro	Cristiano Ronaldo dos Santos Aveiro
GK	David De Gea Quintana	David De Gea Quintana
CB	Giorgio Chiellini	Giorgio Chiellini
ST	-	Kylian Mbappé Lottin

Table 15: Starting Line-Up Squad Selection (For 4-2-3-1 and 4-3-3 Formation)

Data Attribute Explanation

Attacking	
Crossing	The ability to deliver the ball from the flanks into the box.
Finishing	Accuracy and composure when shooting at goal.
Heading accuracy	Skill in directing the ball with the head.
Short passing	Precision in passing the ball over short distances.
Volleys	Technique in striking the ball before it touches the ground.
Skill	
Dribbling	The ability to maneuver the ball past opponents.
Curve	The ability to bend the ball, especially on shots or crosses.
FK Accuracy	Precision in taking free kicks.
Long passing	Precision in passing the ball over long distances.
Ball control	Skill in managing the ball and keeping it under control.
Movement	
Acceleration	How quickly a player can reach top speed.
Sprint speed	The player's top running speed.
Agility	How well a player can start, stop, and move in different directions.
Reactions	How quickly a player responds to situations on the field.
Balance	The ability to stay upright and stable when moving or under pressure.
Power	
Shot power	The force behind shots at goal.
Jumping	The ability to leap high, important for heading and goalkeeping.
Stamina	Endurance and the ability to sustain effort throughout the match.
Strength	Physical power to shield the ball and win physical duels.
Long shots	The ability to accurately shoot from outside the penalty area.
Mentality	
Aggression	The degree of assertiveness in play.
Interceptions	The ability to read the game and cut out opposition passes.
Att. Position	Instinct in taking up positions to create scoring opportunities.
Vision	The ability to see and execute plays that others might not.
Penalties	Skill in taking penalty kicks.
Composure	The ability to stay calm and make good decisions under pressure.
Defending	
Defensive awareness	Understanding of defensive positioning and tactics.
Standing tackle	Ability to dispossess opponents while staying on feet.
Sliding tackle	Skill in taking the ball away with a slide on the ground.
Goalkeeping	
GK Diving	The keeper's ability to dive and save shots.
GK Handling	How well the goalkeeper catches and holds the ball.
GK Kicking	The keeper's skill in playing the ball with their feet.
GK Positioning	The goalkeeper's sense of the best position to be in to make saves.
GK Reflexes	The goalkeeper's reaction speed in making sudden saves.
Special	
Potential	A measure of how good a player could become.
International reputation	The player's prestige on the international stage.

Table 16: Player Abilities and Attributes

Formation Requirement

Position	Skills Required
Defensive Midfielders	Defending, Mentality (Interceptions), Passing, Mentality (Aggression), Physic, Power (Stamina), Mentality (Positioning)
Striker	Shooting, Attacking (Finishing), Pace, Mentality (Composure), Attacking (Heading Accuracy), Skill (Ball Control), Movement (Reactions)
Attacking Midfielders/Wingers	Dribbling, Attacking (Crossing), Skill (Curve), Attacking (Short Passing), Mentality (Vision), Movement (Agility)
Defenders	Defending (Marking Awareness), Power (Strength), Mentality (Composure), Movement (Reactions), Defending (Standing Tackle), Defending (Sliding Tackle)
Goalkeeper	Goalkeeping (Diving), Goalkeeping (Handling), Goalkeeping (Positioning), Goalkeeping (Reflexes), Goalkeeping (Kicking)

Table 17: Skills Required for Different Positions in a 4-2-3-1 Formation

Position	Skills Required
Midfielders	Passing, Skill (Ball Control), Mentality (Vision), Mentality (Positioning), Mentality (Interceptions), Physic
Wingers	Pace, Dribbling, Attacking (Crossing), Shooting, Movement (Reactions), Movement (Agility)
Striker	Shooting, Attacking (Finishing), Attacking (Heading Accuracy), Skill (Ball Control), Mentality (Positioning)
Defenders	Defending, Power (Strength), Movement (Reactions), Defending (Standing Tackle), Defending (Sliding Tackle)
Goalkeeper	Goalkeeping (Diving), Goalkeeping (Handling), Goalkeeping (Positioning), Goalkeeping (Reflexes), Goalkeeping (Kicking)

Table 18: Skills Required for Different Positions in a 4-3-3 Formation

Player Positions Explained

Figure 1: Football Positions

- **GK (Goalkeeper)**: Defends the goal, only player allowed to handle the ball within the penalty area.

- **SW (Sweeper)**: A defensive player who "sweeps up" the ball if the opposition breaks through the defensive line.

- **CB (Center Back)**: Central defenders who stop opposing players, especially strikers, from scoring.

- **LCB/RBC (Left/Right Center Back)**: Play to the left/right of the central defender and help in defending against attacks from the wings.

- **LB/RB (Left/Right Back)**: Fullbacks who defend the left and right flanks and can also provide width in attack.

- **LWB/RWB (Left/Right Wing Back)**: Wide defenders/midfielders who cover lots of ground by supporting both defense and attack on the flanks.

- **CDM (Central Defensive Midfielder)**: Shields the defense by breaking up opposition attacks and starting their team's play.

- **CM (Center Midfielder)**: Central player who connects defense and attack, involved in both passing and tackling.

- **LCM/RCM (Left/Right Center Midfielder)**: Play on the sides of the central midfielder and support both defense and offense.

- **LM/RM (Left/Right Midfielder)**: Wide players tasked with creating chances from the sides and supporting both attack and defense.

- **CAM (Central Attacking Midfielder)**: Creates scoring opportunities, often plays behind the strikers.

- **LW/RW (Left/Right Winger)**: Attack down the left/right flanks, often providing crosses into the box.

- **CF (Center Forward)**: A versatile forward who can both score and create opportunities for other players.

- **SS (Second Striker)**: Plays behind the main striker and supports the attack, often with freedom to roam.

- **ST (Striker)**: The team's primary goal-scorer, positioned closest to the opponent's goal.

Position	LS	ST	RS	LW	LF	CF	RF	RW	LAM	CAM	RAM
Index	0	1	2	3	4	5	6	7	8	9	10
Position	LM	LCM	CM	RCM	RM	LWB	LDM	CDM	RDM	RWB	LB
Index	11	12	13	14	15	16	17	18	19	20	21
Position	LCB	CB	RCB	RB	GK						
Index	22	23	24	25	26						

Table 19: Football Team Positions and Index Numbers

2018 World Cup 23 player squad from Gurobi

Goalkeepers (GK): M. Neuer, M. ter Stegen, T. Horn;
Center Backs (CB): M. Hummels (RCB), J. Boateng (RCB), N. Süle (LCB), J. Tah (LCB);
Full Backs: J. Kimmich (RWB), J. Hector (LB), L. Bender (RB);
Central Midfielders (CM): T. Kroos (LCM), L. Goretzka (LCM), S. Khedira; (RCM)
Attacking Midfielders: M. Özil (RAM), T. Müller (CAM), İ. Gündoğan (RAM), J. Draxler (LAM), J. Brandt (LAM);
Wingers: L. Sané (LW), M. Reus (RM), S. Gnabry (LW);
Strikers (ST): T. Werner, M. Kruse.

Figure 2: 4-3-3 Formation

Figure 3: 3-4-3 Formation

Figure 4: 2018 World Cup Germany starting lineup

2021 Euro Cup 26 player squad from Gurobi

Figure 5: 4-3-3 Formation

Figure 6: 3-4-3 Formation

Figure 7: 2021 Euro Cup Germany starting lineup

Goalkeepers (GK): M. Neuer, B. Leno, K. Trapp;
Center Backs (CB): R. Koch, M. Hummels, J. Boateng, M. Ginter, A. Rüdiger, N. Süle;
Full Backs: R. Gosens, M. Arnold, L. Klostermann;
Central Midfielders (CM): J. Kimmich, T. Kroos, L. Goretzka, İ. Gündoğan, F. Neuhaus
Attacking Midfielders: T. Müller, K. Volland, J. Musiala, K. Havertz J. Brandt;

Wingers: S. Gnabry;
Strikers (ST): T. Werner, L. Sané, E. Can.

Pre-Processing Steps

Manchester City team

1. Dummy encoding roles/positions: The player positions column contains multiple positions separated by commas. We used str.split() to separate these into a list of positions per player. Then MultiLabelBinarizer was used to one-hot encode these positions. This transforms the positions into binary indicator columns.

2. Adaptation score calculation: Defined a list of club IDs with similar playing styles to Guardiola. Checked if each player was at one of those clubs before the season. Created a column with counts of previous clubs. Assigned a score of 1 if they played for a similar club before, 0.5 otherwise.

3. Separating work rate: Split the work rate string on '/' into two new columns; Mapped the text ratings like 'Low' to numeric values from 1-3.

4. Other feature engineering: Counted the number of positions that could be played per player; Extracted from text data injury-prone players and dummified; Flagged players with contracts expiring soon.

5. Other pre-processing: Removed unnecessary ID columns; Filtered to only 2015-09-21 data to ensure it is indeed the planning phase and not after the transfer window phase.

Manchester United team

1. We first used the pandas library to conduct detailed filtering and processing of the player data, which included filling missing values with 0, recording the list of 19 active players into a dictionary, and retrieving their complete information to facilitate subsequent modeling.

2. Additionally, the original data contained a position column that mixed multiple position information into a string, which we parsed and sorted, creating a new dictionary named "player positions" for ease of model design.

Coaching Decision-Making

1. Filtering to German players played in 2018 with an overall rating of 79 or above.

2. Refined the dataset by excluding players affected by injuries, bolstering its reliability.

3. Captured information on preferred foot, overall ratings, and an array of ability values across 27 positions.