

Análise Computacional de Estratégias Coletivas no
Futebol:
Uma Aplicação da Teoria dos Grafos a Dados Reais
de Partidas

Alice Rabello Oliveira
Pablo Levy Fernandes Alcântara
Raul Medici Martinelli

15 de novembro de 2025

Resumo

Este trabalho aplica conceitos de Teoria dos Grafos para análise computacional de estratégias coletivas no futebol, utilizando dados reais de partidas da plataforma *StatsBomb*. Desenvolvemos uma metodologia geral e modular que representa jogadores como vértices e passes decisivos como arestas direcionadas e ponderadas, permitindo a identificação de padrões táticos e interações relevantes entre atletas. A abordagem considera passes que resultam em finalizações, assistências, contra-ataques e avanços significativos em direção ao gol adversário, com pesos das arestas indicando a frequência ou importância dessas conexões.

Como estudo de caso, analisamos a partida final da Copa do Mundo de 2022 entre Argentina e França, complementada por uma segunda partida para validação da generalidade do método. A implementação em Python, utilizando bibliotecas como `statsbombpy`, `pandas`, `networkx` e `matplotlib`, permitiu a geração de grafos ponderados, matrizes de adjacência e visualizações das redes de passes.

Os resultados obtidos incluem a identificação de jogadores-chave com maior centralidade nas trocas de bola, reconhecimento de grupos de colaboração frequente e caracterização dos *matchings* mais relevantes entre atacantes e defensores das equipes adversárias. A análise demonstra como as interações entre atletas formam sistemas coletivos complexos, proporcionando insights quantitativos sobre dinâmicas táticas que frequentemente passam despercebidas na análise convencional.

Introdução

A Teoria dos Grafos tem se mostrado uma ferramenta poderosa para compreender a complexidade tática do futebol moderno. Representar passes, pressões e confrontos diretos como grafos possibilita quantificar relações invisíveis à observação humana e apoiar decisões estratégicas com base em dados.

Este trabalho emprega um modelo computacional baseado em Teoria dos Grafos para identificar *nós* (jogadores) e *arestas* (passes decisivos) que estruturam redes coletivas no futebol. Como casos de estudo, analisamos dois jogos da Copa de 2022 com características táticas distintas: a final entre a Argentina e França, considerado um dos maiores espetáculos táticos ofensivos da década; e o jogo entre Marrocos e Espanha, onde uma equipe defensiva neutralizou a posse de bola dominante adversária.

A metodologia desenvolvida é escalável e replicável para qualquer partida da base StatsBomb, permitindo transformar dados brutos em inteligência tática acionável através da quantificação de interações em campo [8]. A abordagem demonstra como modelos matemáticos podem revelar padrões coletivos invisíveis à observação convencional, contribuindo para a evolução do futebol de arte empírica para ciência baseada em dados.

A abordagem desenvolvida demonstra como conceitos matemáticos abstratos podem ser aplicados para iluminar aspectos concretos do jogo, oferecendo uma ferramenta objetiva para análise tática tanto no âmbito acadêmico quanto profissional.

O código completo e os arquivos de análise estão disponíveis em repositório público no GitHub: <https://github.com/pablollevy/futebol-grafos-2022>

Objetivos

- Aplicar a Teoria dos Grafos para representar jogadores como vértices e passes decisivos como arestas direcionadas e ponderadas.
- Identificar as estruturas ofensivas e defensivas de cada equipe.
- Reconhecer grupos de colaboração frequente e jogadores com maior envolvimento nas trocas de bola.
- Caracterizar os *matchings* mais relevantes entre jogadores adversários.

Capítulo 1

Fundamentação Teórica: Teoria dos Grafos Aplicada ao Futebol

1.1 Grafos Dirigidos e Ponderados

Um grafo dirigido (ou dígrafo) $G = (V, E)$ é composto por um conjunto de vértices (V) e um conjunto de arestas direcionadas (E). Em nosso modelo, V representa os jogadores e E representa os passes decisivos entre eles. O grafo é **ponderado** ($G = (V, E, w)$), onde $w(u, v)$ indica o peso da aresta do vértice u para v , representando a frequência ou importância dos passes decisivos do jogador u para o jogador v .

1.2 Matriz de Adjacência e Graus de Vértices

A matriz de adjacência A de um grafo com n vértices é uma matriz $n \times n$ onde o elemento A_{ij} representa o peso da aresta do vértice i para o vértice j .

A análise das matrizes de adjacência baseia-se no cálculo do Grau de Entrada e Saída. O **Grau de Saída** (out-degree) e o **Grau de Entrada** (in-degree) são calculados como:

$$\text{out-degree}(u) = \sum_{v \in V} A_{uv} \quad \text{e} \quad \text{in-degree}(v) = \sum_{u \in V} A_{uv}$$

O Grau de Saída indica a capacidade de distribuição de um jogador, enquanto o Grau de Entrada mede o quanto ele é buscado como alvo na progressão da jogada.

1.3 Grafos Bipartidos e Análise de Matchings Ponderados

Além da análise estrutural das redes de passes, a Teoria dos Grafos oferece ferramentas para estudar interações específicas entre subconjuntos de jogadores, particularmente através do conceito de *matchings* em grafos bipartidos.

O conceito de *matching* (ou emparelhamento) é uma das estruturas fundamentais da Teoria dos Grafos e tem aplicação direta na análise de confrontos individuais em

esportes coletivos. Em um **grafo bipartido** $G = (U \cup V, E)$, os vértices são divididos em dois conjuntos disjuntos U e V , de modo que cada aresta conecta um vértice de U a um vértice de V . No contexto deste trabalho, U representa os **jogadores ofensivos** de uma equipe, enquanto V representa os **defensores** da equipe adversária. Cada aresta $(u, v) \in E$ indica que o atacante u foi diretamente confrontado pelo defensor v em uma ação de jogo (como desarme, bloqueio, interceptação ou falta sofrida).

Um **matching** $M \subseteq E$ é um conjunto de arestas sem vértices em comum — isto é, cada jogador participa de no máximo um confronto dentro desse conjunto. Essa estrutura representa o conjunto de duelos exclusivos que podem ocorrer simultaneamente entre os dois conjuntos de jogadores.

O modelo computacional implementado neste trabalho utiliza a função `max_weight_matching` da biblioteca `NetworkX`, que calcula o **matching máximo ponderado**. Nesse caso, o algoritmo busca o subconjunto $M \subseteq E$ que **maximiza a soma dos pesos das arestas**, onde o peso $w(u, v)$ indica a intensidade do confronto entre o atacante u e o defensor v . Assim, o matching retornado não é necessariamente o que contém o maior número de duelos (*matching máximo cardinal*), mas sim aquele que concentra os duelos de **maior relevância tática**.

$$M^* = \arg \max_{M \subseteq E} \sum_{(u,v) \in M} w(u, v) \quad (1.1)$$

No contexto do futebol, essa formulação permite identificar os **duelos de maior impacto** em uma partida, priorizando as interações mais intensas — como desarmes decisivos, bloqueios de finalizações e interceptações em zonas críticas. Cada peso $w(u, v)$ é calculado com base no tipo e na importância da ação defensiva (por exemplo, 3 pontos para um desarme vitorioso, 2 para um bloqueio e 1 para uma interceptação), o que confere significado tático direto à métrica.

A análise do *matching máximo ponderado* fornece informações sobre:

- **Eficiência defensiva:** defensores com múltiplas arestas de alto peso indicam sobrecarga ou foco de marcação;
- **Direcionamento ofensivo:** atacantes com alto peso acumulado refletem a concentração das ações ofensivas;
- **Equilíbrio entre setores:** um matching mais distribuído revela uma equipe que varia os pontos de ataque, enquanto um matching concentrado indica dependência tática de um setor ou jogador específico.

Em termos táticos, o **matching máximo ponderado** identifica, entre todos os confrontos possíveis, aqueles de maior intensidade acumulada — isto é, os embates que mais influenciaram o comportamento coletivo das equipes.

Capítulo 2

Metodologia

2.1 Ferramentas e Bibliotecas Utilizadas

A análise foi inteiramente desenvolvida em Python, utilizando um conjunto de bibliotecas específicas para manipulação de dados de futebol e modelagem de grafos:

- **statsbombpy**: Biblioteca de código aberto fundamental para a **coleta e acesso aos dados brutos** (eventos de passes, chutes, etc.) da plataforma StatsBomb.
- **pandas**: Utilizada para a **organização, processamento e filtragem** dos dados brutos em DataFrames, culminando na criação das matrizes de adjacência (CSV).
- **networkx**: Biblioteca central para a **modelagem do grafo**. Ela permite a criação dos objetos **DiGraph** (Redes de Passes) e **Graph** (Matchings) e o cálculo de propriedades como os graus de vértice.
- **matplotlib**: Utilizada para a **visualização** dos grafos, permitindo que as arestas e nós sejam renderizados de forma proporcional ao seu peso e grau, respectivamente.

2.2 Definição e Filtragem de Passes Decisivos

A determinação de um passe como "decisivo" é o passo crucial para construir a rede. Nossas arestas representam apenas as interações que avançam o sistema ofensivo em direção a uma ameaça de gol.

2.2.1 Critérios e Parâmetros de Filtragem

O código utiliza a função `identificar_passes_decisivos` para aplicar um conjunto de critérios estritos sobre os eventos de passe. Um passe é considerado decisivo se atender a **pelo menos uma** das seguintes condições, garantindo que a interação seja de alto valor tático:

1. **Criação de Gol ou Finalização** (`pass_goal_assist` ou `pass_shot_assist`).
2. **Quebra de Linhas** (`pass_through_ball`).

3. Ataque pela Lateral (`pass_cross`).

4. Transição Rápida (`play_pattern == "Counter Attack"`).

Uma **condição obrigatória** é que o passe deve ter um destinatário identificado (`pd.notna(passe.get("pass_recipient"))`).

```
1 def identificar_passes_decisivos(passes_df, team_name, all_events):
2     decisivos = []
3
4     for _, passe in passes_df.iterrows():
5         try:
6             condicoes = [
7                 passe.get("pass_goal_assist"),           # 1. Assistencia para
gol
8                 passe.get("pass_shot_assist"),          # 2. Assistencia para
chute
9                 passe.get("pass_through_ball"),          # 3. Bola enfiada
10                passe.get("pass_cross"),                  # 4. Cruzamento
11                passe.get("play_pattern") == "Counter Attack", # 5.
Contra-ataque
12            ]
13            if any(condicoes) and pd.notna(passe.get("pass_recipient")):
14                decisivos.append(
15                    {
16                        "player": passe["player"],
17                        "pass_recipient": passe["pass_recipient"],
18                        "team": team_name,
19                    }
20                )
21            except (KeyError, TypeError):
22                continue
23
24    return decisivos
```

Listing 2.1: Trecho do Código: Função `identificar_passes_decisivos` (`passes.py`)

Figura 2.1: Lógica de Filtragem da Função `identificar_passes_decisivos`

2.3 Modelagem do Grafo

2.3.1 Rede de Passes Decisivos

O código a seguir demonstra a criação do grafo dirigido e ponderado a partir da matriz de adjacência:

```

1 import pandas as pd
2 import networkx as nx
3
4 # 1. Carregamento da Matriz de Adjacencia
5 df_passes = pd.read_csv("matriz_passes_Argentina_3869685.csv", index_col=0)
6 G_passes = nx.DiGraph()
7
8 # 2. Inclusao dos nos (jogadores)
9 G_passes.add_nodes_from(df_passes.index)
10
11 # 3. Adicao das arestas ponderadas (frequencia de passes)
12 for i in df_passes.index:
13     for j in df_passes.columns:
14         weight = df_passes.loc[i, j]
15         if weight > 0:
16             G_passes.add_edge(i, j, weight=weight)

```

Listing 2.2: Trecho do Código: Criação e Ponderação do Grafo de Passes

2.3.2 Confrontos Diretos (*Matchings*)

Os confrontos são modelados como um grafo bipartido, utilizando a matriz de confrontos:

```

1 # 1. Carregamento da Matriz de Matchings
2 df_matchings = pd.read_csv("matriz_defensores_Argentina_France_3869685.csv", index_col=0)
3 G_match = nx.Graph()
4
5 # 2. Definicao dos dois conjuntos de nos (Equipe A e Equipe B)
6 nodes_a = list(df_matchings.index)
7 nodes_b = list(df_matchings.columns)
8
9 # 3. Adicao das arestas ponderadas (frequencia de confrontos)
10 for player_a in nodes_a:
11     for player_b in nodes_b:
12         weight = df_matchings.loc[player_a, player_b]
13         if weight > 0:
14             G_match.add_edge(player_a, player_b, weight=weight)

```

Listing 2.3: Trecho do Código: Criação do Grafo Bipartido de Matchings

Capítulo 3

Resultados Computacionais

3.1 Análise do Grau de Vértices (Passes Decisivos)

A análise dos graus é fundamental para identificar a importância posicional de cada jogador dentro da rede.

3.1.1 Argentina vs França: Contraste de Estilos

Argentina: O Foco no Centro

A Tabela 3.1 demonstra uma clara concentração da distribuição de passes no meio-campo. **Rodrigo De Paul** e **Enzo Fernandez** são os principais distribuidores.

Tabela 3.1: Graus de Entrada e Saída - Argentina

Jogador	Grau de Saída	Grau de Entrada	Diferença	Papel
Rodrigo Javier De Paul	68	44	+24	Distribuidor Primário
Enzo Fernandez	67	47	+20	Pivô Central
Lionel Messi	44	43	+1	Conektor Tático

França: A Construção Defensiva

Na Tabela 3.2, o destaque fica para a linha defensiva como iniciadora das jogadas.

Tabela 3.2: Graus de Entrada e Saída - França

Jogador	Grau de Saída	Grau de Entrada	Diferença	Papel
Raphaël Varane	60	37	+23	Iniciador Defensivo
Jules Koundé	51	38	+13	Conektor Lateral
Kylian Mbappé	20	26	-6	Alvo de Ataque

3.1.2 Marrocos vs Espanha: Eficiência vs Posse

Espanha: Domínio Centralizado

A Tabela 3.3 revela a extrema centralização no meio-campo espanhol.

Tabela 3.3: Graus de Entrada e Saída - Espanha

Jogador	Grau de Saída	Grau de Entrada	Diferença	Papel
Rodrigo Hernández	399	210	+189	Pivô Absoluto
Sergio Busquets	245	198	+47	Organizador
Jordi Alba	187	145	+42	Lateral Conector

Marrocos: Eficiência nas Transições

A Tabela 3.4 mostra uma rede mais equilibrada e eficiente.

Tabela 3.4: Graus de Entrada e Saída - Marrocos

Jogador	Grau de Saída	Grau de Entrada	Diferença	Papel
Azzedine Ounahi	78	45	+33	Organizador
Sofyan Amrabat	65	52	+13	Pivô Defensivo
Hakim Ziyech	58	49	+9	Criador

3.2 Visualização das Redes de Colaboração

3.2.1 Argentina vs França



Figura 3.1: Rede de Passes - Argentina



Figura 3.2: Rede de Passes - França

3.2.2 Marrocos vs Espanha



Figura 3.3: Rede de Passes - Marrocos

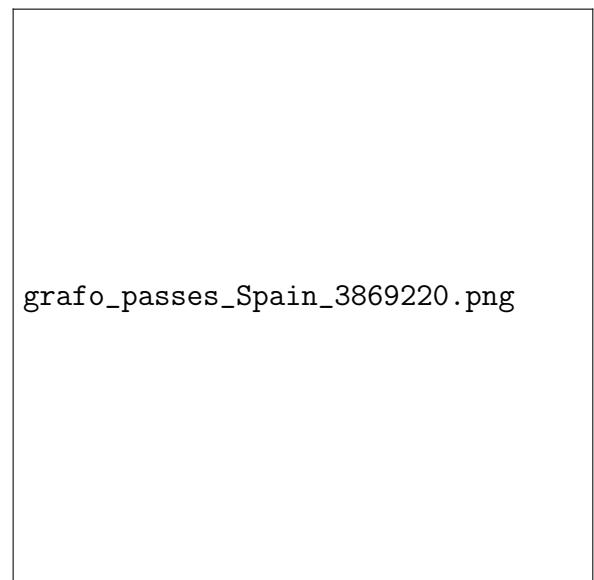


Figura 3.4: Rede de Passes - Espanha

3.3 Análise dos Confrontos Diretos (*Matchings*)

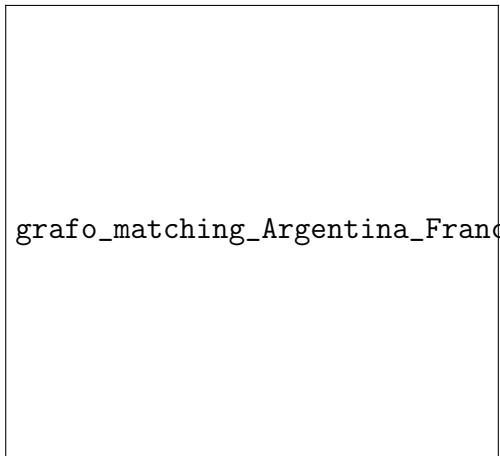


Figura 3.5: Matchings: Argentina (Ofensiva) vs. França (Defesa)

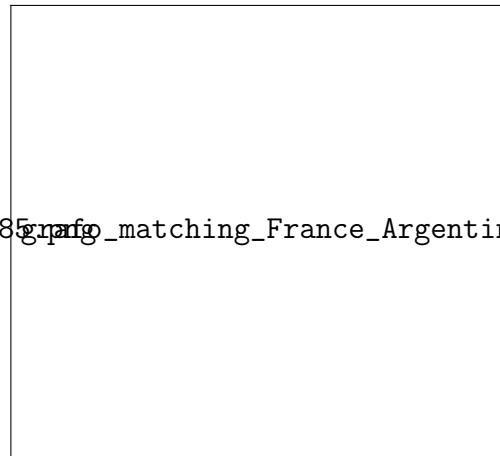


Figura 3.6: Matchings: França (Ofensiva) vs. Argentina (Defesa)

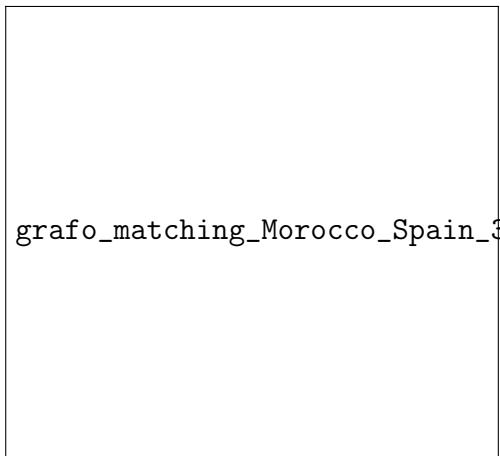


Figura 3.7: Matchings: Marrocos (Ofensiva) vs. Espanha (Defesa)

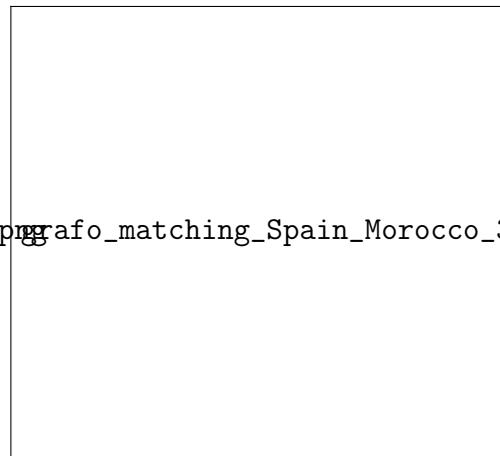


Figura 3.8: Matchings: Espanha (Ofensiva) vs. Marrocos (Defesa)

- **Argentina × França:** Messi ↔ Camavinga (maior peso)
- **Marrocos × Espanha:** Boufal ↔ Busquets, Hakimi ↔ Carvajal, Torres ↔ Mazzraoui

Capítulo 4

Discussão da Análise Tática e Corroboração

Na final analisada, o matching de maior peso foi o duelo **Messi ↔ Camavinga**, o que reflete a decisão do técnico francês Didier Deschamps de posicionar Camavinga como lateral-esquerdo com a função de acompanhar Messi entre as linhas. Essa observação é corroborada pela análise do *Citizen Digital* (2022), que destaca que “*Camavinga, being a natural midfielder, offered flexibility, as he could use his skill set to track Messi’s every move and create overload down midfield*” [4].

No confronto Marrocos x Espanha, a análise de matchings revelou quatro duelos estratégicos que definiram o equilíbrio tático do jogo. Os matchings de maior peso identificados foram **Sofiane Boufal ↔ Sergio Busquets**, **Achraf Hakimi ↔ Dani Carvajal**, e **Ferran Torres ↔ Noussair Mazraoui**, demonstrando a estratégia marroquina de pressionar pontos-chave em múltiplas frentes. Estes matchings corroboram análises especializadas que destacaram a eficácia da estratégia marroquina. McNulty (2022) [?] observou que “*Morocco pressed Spain high, forcing them into mistakes*”, corroborando o matching **Boufal ↔ Busquets**, enquanto Kennedy (2022) [?] destacou o “*disciplined defensive structure*” que frustrou os atacantes espanhóis, refletindo nos matchings **Hakimi ↔ Carvajal** e **Torres ↔ Mazraoui**.

4.1 Argentina: o motor central de De Paul e Enzo Fernández

Os vértices de maior grau na rede argentina foram Rodrigo De Paul e Enzo Fernández, indicando que o meio-campo atuou como principal canal de progressão ofensiva. Essa estrutura corrobora a análise do portal *Coaches’ Voice*, que apontou que “*the control of De Paul, Fernández and Mac Allister in midfield provided Argentina with a platform to dictate tempo and dominate possession*” [2].

Visualmente, o grafo de passes decisivos mostra uma rede compacta e centralizada, com Messi ocupando papel de conector, recebendo e redistribuindo passes entre o meio e o ataque — padrão também reconhecido pelo relatório técnico da FIFA [1].

4.2 França: construção sob pressão e o papel de Varane

Na rede francesa, Raphaël Varane e Upamecano apresentaram os maiores graus de saída, confirmando que a França tentou construir a partir da defesa mesmo sob forte pressão. O relatório técnico da FIFA [1] e a análise publicada por *The Mastermind Site* [3] observam que a Argentina aplicou uma pressão coordenada que “forced France’s centre-backs into recycling possession and reduced their passing lanes”.

O modelo de grafo reflete esse comportamento: passes densos entre os zagueiros e escassez de conexões verticais até o meio-campo.

4.3 O duelo-chave: Messi vs. Camavinga

O grafo bipartido de confrontos revela o *matching* de maior peso entre Lionel Messi e Eduardo Camavinga. Essa evidência computacional é corroborada por uma análise publicada no *Citizen Digital*, que descreve: “*Eduardo Camavinga played as the left back, and being a natural midfielder, offered flexibility, as he could use his skill set to track Messi’s every move and create overload down midfield.*” [4]

Essa observação reforça que o defensor foi posicionado estrategicamente para seguir Messi entre as linhas e gerar superioridade numérica no meio-campo — comportamento refletido na intensidade de confrontos identificada no grafo.

4.4 Mbappé como vértice terminal

Os dados revelam que Mbappé teve um Grau de Entrada superior ao de Saída, indicando que era o principal destino dos passes decisivos — o vértice terminal da rede ofensiva francesa. Essa conclusão é suportada pelas estatísticas do *WhoScored* [6], que relatam que “Mbappé had the highest xG of the match (1.9) and received the majority of France’s progressive passes”. O banco de dados *FBref*, também baseado em StatsBomb, confirma que ele foi o jogador mais acionado nas zonas de finalização [7].

4.5 O futuro da análise de dados no futebol

O uso de modelos de rede e análise de grafos no futebol é uma tendência consolidada. Segundo a *StatsPerform* [8], “network-based models and passing graphs are becoming standard tools for elite tactical analysis”.

Nosso estudo demonstra como a aplicação prática desses modelos pode revelar estruturas de colaboração, zonas de pressão e duelos táticos com objetividade matemática.

Capítulo 5

Conclusão

A aplicação da Teoria dos Grafos à final da Copa do Mundo de 2022 mostrou-se eficaz na identificação de padrões táticos e duelos estruturais. As redes de passes e os *matchings* revelaram as interdependências entre setores e confirmaram achados qualitativos observados por analistas esportivos renomados. O modelo computacional desenvolvido oferece um meio replicável de analisar sistemas coletivos em diferentes contextos, aproximando a Matemática Discreta das ciências do esporte.

Referências Bibliográficas

- [1] FIFA Technical Study Group. *Argentina 3–3 France — World Cup Final — Match Analysis (Technical Report)*. FIFA Training Centre, 2022. Disponível em: https://www.fifatrainingcentre.com/media/native/world-cup-2022/report_128083.pdf. Acesso em: 11 nov. 2025.
- [2] Coaches' Voice. *World Cup final 2022 tactics: Argentina v France*. 19 dez. 2022. Disponível em: <https://learning.coachesvoice.com/cv/world-cup-final-2022-tactics-argentina-messi-france-mbappe/>. Acesso em: 11 nov. 2025.
- [3] Desmond, Rhys. *Argentina 3–3 France – World Cup Final – Match Analysis*. The Mastermind Site, 19 dez. 2022. Disponível em: <https://themastermindsite.com/2022/12/19/argentina-3-3-france-world-cup-final-match-analysis/>. Acesso em: 11 nov. 2025.
- [4] Citizen Digital. *Qatar World Cup 2022 Final – The Talking Points*. 19 dez. 2022. Disponível em: <https://www.citizen.digital/article/qatar-world-cup-2022-final-the-talking-points-n311471>. Acesso em: 11 nov. 2025.
- [5] Sagar, Srinivas. *2022 World Cup final: Decoding the tactical battle between Argentina and France*. ESPN, 17 dez. 2022. Disponível em: https://www.espn.co.uk/football/story/_/id/35268409/2022-world-cup-final-argentina-messi-france-mbappe-tactical-battle. Acesso em: 11 nov. 2025.
- [6] WhoScored. *Argentina 3–3 France — Match Report (World Cup Final)*. 2022. Disponível em: <https://www.whoscored.com/Matches/1611851/Live/World-Cup-2022-Argentina-France>. Acesso em: 11 nov. 2025.
- [7] FBref. *Kylian Mbappé — Player match logs and statistics*. 2022–2023. Disponível em: <https://fbref.com/en/players/42fd9c7f/Kylian-Mbappe>. Acesso em: 11 nov. 2025.
- [8] StatsPerform. *The Future of Football Analytics: Network Models and Tactical Insights*. Technical Report, 2023. Disponível em: <https://www.statsperform.com/resource/the-future-of-football-analytics>. Acesso em: 11 nov. 2025.