

Um *Framework* Baseado em Otimização e Séries Temporais para Composição de Equipes no Cartola-FC

Letícia Maia Távora

Programa de Engenharia de Sistemas e Computação
Universidade Federal do Rio de Janeiro
leticiamt@cos.ufrj.br

Israel Mendonça

Faculty of Advanced Science and Technology
Kumamoto University
israel@cs.kumamoto-u.ac.jp

Glaucio Fiorott Amorim

Departamento de Estatística
Centro Federal de Educação Tecnológica Celso Suckow da Fonseca
glaucio.amorim@cefet-rj.br

Pedro Henrique González

Programa de Engenharia de Sistemas e Computação
Universidade Federal do Rio de Janeiro
pegonzalez@cos.ufrj.br

RESUMO

Com a crescente popularidade dos *Fantasy Games*, como o Cartola FC, surge a necessidade de estratégias eficientes para a composição de equipes virtuais. Neste trabalho, adotamos métodos de previsão e análise de dados históricos do Campeonato Brasileiro com o objetivo de maximizar a pontuação das equipes, reduzindo custos virtuais e evitando escolhas que comprometam o desempenho. Propomos a utilização de diferentes métodos preditivos para antecipar o desempenho futuro dos jogadores, aplicando técnicas de otimização para selecionar a melhor formação possível, considerando restrições de custo e posição. Os resultados mostraram que o modelo ARMA apresentou maior precisão na previsão de desempenho dos jogadores, em comparação a outros métodos de previsão. Esta abordagem sistemática permite aos usuários tomar decisões estratégicas mais informadas, maximizando o desempenho de suas equipes no Cartola FC.

PALAVRAS CHAVE. Otimização Combinatória, Séries Temporais, Cartola FC.

OA – Outras aplicações em PO, IC – Inteligência Computacional

ABSTRACT

With the growing popularity of *Fantasy Games* such as Cartola FC, there is a need for efficient strategies for composing virtual teams. In this work, we adopt methods of prediction and analysis of historical data from the Brazilian Championship with the aim of maximizing team scores, reducing virtual costs and avoiding choices that compromise performance. We propose the use of different predictive methods to anticipate the future performance of players, applying optimization techniques to select the best possible formation, considering cost and position constraints. The results showed that the ARMA model was more accurate in predicting player performance than other prediction methods. This systematic approach allows users to make more informed strategic decisions, maximizing the performance of their teams on Cartola FC.

KEYWORDS. Combinatorial Optimization, Time Series, Cartola FC.

OA – Other applications in OR, IC – Computational Intelligence

1. Introdução

O Cartola FC, um jogo virtual de fantasia baseado no futebol, vem atraindo uma legião crescente de entusiastas, representando uma interseção fascinante entre esporte, análise de dados e entretenimento digital [GloboEsporte.com, 2017]. Os usuários discutem estratégias, compartilham *insights* e criam conteúdos relacionados, aumentando a interação e o engajamento em torno do futebol [Casagrande et al., 2021]. Essa interação não só enriquece a experiência do usuário, mas também contribui para a construção de uma base de fãs mais sólida e fiel. Além de seu valor como entretenimento, o Cartola FC também apresenta um potencial financeiro significativo, fomentando uma comunidade ativa e engajada. A movimentação financeira em torno do Cartola FC é multifacetada e substancial. A plataforma atrai investimentos de várias fontes, incluindo patrocínios, publicidade e a venda de dados. Empresas têm reconhecido o potencial do Cartola FC como um meio eficaz de se conectar com um público altamente segmentado e engajado [Farquhar e Meeds, 2007]. Isso abre um leque de oportunidades para empresas e entidades que desejam se envolver com essa base de fãs apaixonada por futebol e tecnologia.

Ao incorporar técnicas avançadas de análise de dados, este estudo demonstra como o Cartola FC pode ser utilizado de forma analítica para aprimorar a experiência do torcedor e impulsionar possíveis retornos financeiros. Ao examinar padrões de desempenho de jogadores de forma estratégica, buscamos oferecer *insights* valiosos e viáveis. Sendo assim, este trabalho visa o desenvolvimento de um *Framework* que propõe o uso dos resultados dos modelos preditivos para antecipar as pontuações dos jogadores em rodadas futuras e maximizar a pontuação das equipes dos usuários do Cartola FC, se baseando nos resultados dos modelos propostos. Para isso, não só analisamos o desempenho individual dos jogadores, mas também consideramos suas pontuações combinadas podem ser otimizadas para alcançar a melhor escalação com o maior retorno possível.

Adiante, este trabalho está dividido da seguinte forma: a Seção 2 apresenta os trabalhos relacionados ao nosso estudo, assim como a forma que problemas semelhantes estão sendo avaliados e resolvidos na literatura. Na Seção 3, é apresentada a definição do problema do presente estudo. Na Seção 4 se encontra a metodologia utilizada para a resolução do problema. Na sequência, na Seção 5, são abordados os experimentos computacionais realizados e por fim, na Seção 6, são apresentadas as conclusões e sugestões para trabalhos futuros.

2. Trabalhos Relacionados

O uso crescente de análises aplicadas em todos os estratos do esporte [Singh, 2020], sobretudo no futebol, tem se tornado um aliado essencial na tomada de decisões estratégicas. No entanto, a mera acumulação de informações não é suficiente para obter *insights* profundos e tomar decisões eficazes no Cartola FC. Nesse contexto, a necessidade de revisão da literatura torna-se imperativa, uma vez que é preciso aplicar técnicas sofisticadas de análise de dados, como métodos preditivos, para obter respostas e fundamentações sólidas nas estratégias a serem escolhidas.

Sendo assim, a implementação de algoritmos preditivos no futebol é fundamental para este estudo. Embora não sejam focados diretamente em *Fantasy Game*, estudos como o de [Horvat e Job, 2020] permitem aos participantes prever o desempenho dos jogadores com base em dados históricos e estatísticos, oferecendo ferramentas valiosas para este contexto. A utilização de dados históricos ajuda a construir modelos que capturam padrões e tendências no desempenho dos jogadores, proporcionando uma base sólida para previsões acuradas.

O escopo então é ampliado a fim de oferecer uma visão mais detalhada sobre como os modelos de aprendizado de máquina são aplicados para prever pontos em ligas de *Fantasy Football* [Bangdiwala et al., 2022]. Essas metodologias específicas e seus resultados práticos permitem que os participantes do Cartola FC compreendam como os algoritmos preditivos podem ser utilizados para estimar os pontos que os jogadores podem alcançar, influenciando diretamente as decisões de seleção de equipes.

Além disso, a exploração da modelagem de otimização em jogos de esporte fantasia, como discutido por [Beliën et al., 2017], fornece uma abordagem sistemática para a seleção de jogadores. Examinando diferentes estratégias para compor equipes vencedoras, considera variáveis cruciais que influenciam a eficácia das equipes montadas pelos participantes do Cartola FC. As estratégias de otimização contidas neste estudo, ajudam a balancear a composição do time, maximizando a pontuação potencial com base nas restrições e objetivos do jogo.

No mesmo sentido, um estudo abrangente sobre previsão em jogos de *Fantasy Game* é apresentado por [Robinson, 2020]. Essa pesquisa oferece uma visão mais ampla das estratégias de previsão que podem ser adaptadas e aplicadas ao contexto do Cartola FC. A abordagem detalhada deste estudo enriquece o repertório de métodos de previsão disponíveis para os participantes, permitindo a adaptação de técnicas avançadas para melhorar suas estratégias de seleção de jogadores.

Por fim, [Cortinhas et al., 2023] propõem a aplicação de técnicas de otimização combinatória multiobjetivo em dados do campeonato de 2019 do Cartola FC, onde é visado desenvolver uma ferramenta que otimize o uso dos recursos disponíveis, auxiliando os usuários a escalar um time ideal a cada rodada.

Uma observação importante nos estudos analisados é a forma como a previsão e análise de desempenho dos jogadores são abordadas. Em geral, quando essa informação está disponível, são utilizados dados históricos e estatísticos para estimar a performance dos jogadores, como na análise de desempenho baseada em estatísticas históricas [Horvat e Job, 2020] ou em modelos preditivos como em [Bangdiwala et al., 2022]. No entanto, um aspecto crítico frequentemente subestimado é o custo de um resultado otimizado.

Embora o estudo de [Cortinhas et al., 2023] aborde de maneira semelhante à nossa proposta sobre otimização combinatória, ao empregar dados reais do campeonato, há limitações em termos da utilização de diferentes esquemas táticos e aplicação de diferentes métodos de previsões à performance dos jogadores e à formação das equipes, seguindo o mesmo modelo de otimização deste estudo porém diferenciando-se no contexto de análises preditivas. Essas limitações incluem a necessidade de algoritmos mais robustos que possam oferecer previsões mais precisas em cenários dinâmicos e adaptação a mudanças imprevistas nas condições dos jogadores ou nas estratégias das equipes. A Tabela 1 sintetiza os trabalhos relacionados, fazendo uma comparação entre eles e apresentando o diferencial do trabalho que desenvolvemos.

Artigos	Esquema Tático Dinâmico	Custo analisado	Análises preditivas
[Budgen e Brereton, 2006]			X
[Beliën et al., 2017]			X
[Silva, 2019]		X	
[Horvat e Job, 2020]		X	X
[Robinson, 2020]		X	X
[Bangdiwala et al., 2022]			X
[Cortinhas et al., 2023]		X	
Nosso estudo	X	X	X

Tendo em vista o exposto, este estudo busca tratar essas limitações ao desenvolver uma abordagem mais robusta e adaptável para a previsão e otimização no Cartola FC. Ao integrar técnicas de previsão em conjunto com modelos de otimização combinatória, visamos melhorar a precisão das previsões e a eficácia das estratégias de seleção de jogadores.

3. Metodologia Proposta

No contexto da análise de composição de equipes em *Fantasy Games*, a análise de séries temporais pode ser usada para prever o desempenho de um jogador ao longo do campeonato com base em seu histórico de resultados, estatísticas de jogadores e outros fatores relevantes, como o custo e pontuação individual.

Nesse sentido, o problema que buscamos resolver é a composição de equipes no Cartola FC, um jogo de fantasia baseado no futebol brasileiro. Neste jogo, os participantes devem escalar uma equipe de jogadores reais e ganham pontos com base no desempenho desses jogadores nas partidas reais do campeonato. O desafio é selecionar uma equipe que maximize a pontuação total, considerando as restrições de orçamento e as regras do jogo. Nossa abordagem visa fornecer uma solução para esse problema, utilizando métodos de previsão e análise de dados para antecipar o desempenho dos jogadores e técnicas de otimização para selecionar a melhor combinação de jogadores dentro das restrições fornecidas. A dinâmica da escalação do elenco nos *fantasy games* pode ser modelada como um problema de otimização mono-objetivo. Sendo, por conseguinte, um problema com finalidade de maximizar a performance do time escalado no que se refere a escolha dos jogadores mais habilidosos por posição tendo como restrição o dinheiro disponível.

O dinheiro fictício utilizado no jogo denomina-se *Cartoleta*. No início do campeonato é disponibilizada uma quantidade de 100 Cartoletas para montar o seu time. Este valor é decrementado da soma dos custos de todos os jogadores escalados em cada rodada. Após o fim da rodada, o time escalado é vendido (com seu valor modificado em função da performance na rodada atual) e o valor de Cartoletas obtido na venda dos jogadores é somado ao valor de Cartoletas que, porventura, não tivessem sido utilizadas na rodada anterior.

O problema proposto pode ser definido matematicamente como:

$$Opt = \{s \in S \mid \forall \bar{s} \in S, \text{ tal que } \bar{s} \neq s, f(\bar{s}) \leq f(s)\} \quad (1)$$

Em (1) é descrito que para todo s , vetores soluções, pertencente a S , não existe solução $\bar{s}$, em S , tal que $f(\bar{s}) > f(s)$.

Nesta equação, s é uma equipe composta por jogadores, e $f(s)$ é a função objetivo que avalia a pontuação esperada dessa equipe com base nos desempenhos previstos dos jogadores. O problema de otimização visa encontrar a equipe $s \in S$ que maximize $f(s)$, ou seja, que maximize a pontuação total prevista dentro das restrições de orçamento (Cartoletas) e composição de equipe.

Essa abordagem envolve a utilização de métodos de previsão para estimar o desempenho de cada jogador em uma rodada futura e técnicas de otimização para selecionar a combinação de jogadores que maximize a pontuação do time, respeitando o limite de Cartoletas disponível. Essa equação formaliza o objetivo do problema de otimização, que é selecionar a melhor combinação de jogadores s que maximiza a pontuação total $f(s)$. Ela é necessária porque descreve matematicamente o critério de escolha da melhor solução s , deixando claro que não há outra solução $\bar{s}$ que forneça uma pontuação $f(\bar{s})$ maior do que $f(s)$.

Visando a resolução do problema, o restante dessa seção é dedicada a apresentação da metodologia proposta para a resolução do problema definido em (1). Neste estudo é proposto um *Framework* para fornecer previsões precisas e otimizadas do desempenho dos jogadores, seguindo um fluxo cíclico, segmentado pela aplicação de métodos de previsão, otimização da seleção dos jogadores e atualização contínua dos recursos financeiros (cartoletas). A estrutura deste *Framework* permite uma análise dinâmica e adaptativa, garantindo que as decisões sejam sempre informadas pelos dados mais recentes.

Para obtermos um desempenho satisfatório no Cartola FC, desenvolvemos um *Framework* abrangente que integra métodos de previsão, processos de otimização e atualizações dinâmicas. O

fluxo operacional do *Framework* é contínuo e cíclico, garantindo que as decisões sejam sempre baseadas nos dados mais recentes e precisos. Com os dados preparados, avaliamos três métodos de previsão distintos: Médias Móveis, Alisamento Exponencial e ARMA. Essas previsões são então usadas no processo de otimização, onde selecionamos a melhor combinação de jogadores para maximizar o potencial de pontos da equipe, respeitando as restrições de orçamento e as regras do jogo. Utilizamos algoritmos de otimização para explorar todas as possíveis combinações de jogadores possíveis

O *Framework* estruturado na Figura 1 oferece uma abordagem sistemática para a definição de melhores estratégias a serem tomadas. A integração de métodos de previsão, otimização discreta e atualização dinâmica de recursos assegura que as estratégias de seleção de jogadores sejam constantemente refinadas e ajustadas, proporcionando uma vantagem competitiva significativa aos participantes do *Fantasy Game*.

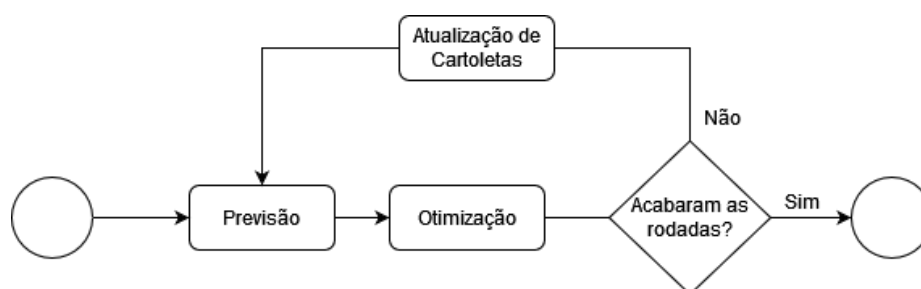


Figura 1: Fluxograma do *Framework* proposto

3.1. Métodos de Previsão

Um dos objetivos deste estudo é prever o desempenho dos jogadores no *Fantasy Game* Cartola FC, utilizando técnicas quantitativas de previsão. A precisão dessas previsões pode proporcionar uma vantagem significativa aos jogadores do Cartola FC, permitindo uma tomada de decisão mais informada na escolha dos jogadores. Para alcançar esse objetivo, foram selecionados três métodos de previsão para serem avaliados após a verificação de estacionariedade dos dados: Médias Móveis, Alisamento Exponencial e ARMA (*AutoRegressive Moving Average*). O fluxo de avaliação dos métodos de previsão seguirá como sugerido na Figura 2

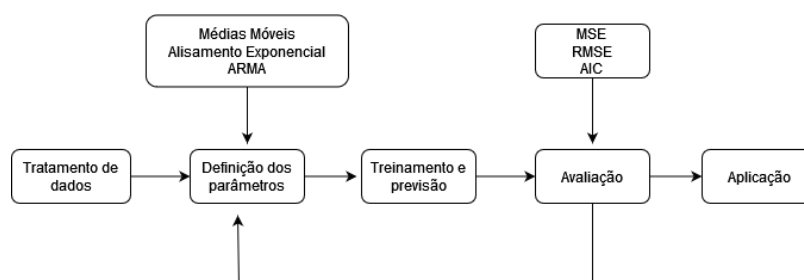


Figura 2: Fluxo de previsão

A técnica de Médias Móveis Simples (MMS) é um método de previsão de curto prazo que se baseia na média de um número determinado de observações anteriores de uma série temporal. Essa técnica visa suavizar as flutuações aleatórias dos dados e revelar a tendência subjacente, permitindo prever os valores futuros com base na dinâmica recente da série. O seu cálculo é dado da seguinte forma:

$$MMS_t = \frac{\sum_{i=t-n}^t x_i}{n} \quad (2)$$

Onde:

MMS_t é a média móvel simples no período t , x_i é a pontuação do jogador na rodada t , que é um dado histórico utilizado para calcular a média móvel e por fim, n sendo o número de rodadas consideradas para o cálculo da média móvel.

Por sua vez, a Alisamento Exponencial Simples (AES) é uma técnica de previsão que se baseia no princípio de dar mais peso aos dados mais recentes. Essa técnica permite atribuir diferentes pesos às observações passadas, dando mais importância aos dados mais recentes, sendo particularmente útil para capturar mudanças rápidas no desempenho de um jogador e produzir previsões mais precisas. Para sua aplicação é usada o seguinte cálculo:

$$Nível_t = \alpha x_t + (1 - \alpha)Nível_{t-1} \quad (3)$$

$$Tendência_t = \beta(Nível_t - Nível_{t-1}) + (1 - \beta)Tendência_{t-1} \quad (4)$$

$$Previsão_t = Nível_t + Tendência_t \quad (5)$$

sendo $Nível_t$ o nível estimado no período t , $Tendência_t$ é a tendência estimada no período t , $Previsão_t$ é a previsão para o período t , α e β são parâmetros de suavização de nível e tendência, respectivamente, e possuem valor entre 0 e 1.

Por fim, o modelo ARMA é amplamente utilizado na análise de séries temporais devido à sua capacidade de modelar componentes autorregressivos e de médias móveis, permitindo uma compreensão profunda das dependências temporais nos dados históricos de desempenho dos jogadores e combina duas partes: a parte autorregressiva (AR) e a parte de média móvel (MA). A parte autorregressiva (AR) do modelo utiliza uma regressão dos valores passados da série para prever os valores futuros. A ordem da parte autorregressiva, denotada por p , indica o número de termos autorregressivos utilizados no modelo. A parte de média móvel (MA) do modelo utiliza os erros de previsão passados (também conhecidos como resíduos) para ajustar as previsões. A ordem da parte de média móvel, denotada por q , indica o número de termos de média móvel utilizados no modelo.

O modelo ARMA(p, q) pode ser formalmente descrito pela seguinte equação:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t$$

Onde y_t é o valor da série no período t , $\phi_1, \phi_2, \dots, \phi_p$ são os coeficientes da parte autorregressiva (AR), $\theta_1, \theta_2, \dots, \theta_q$ são os coeficientes da parte de média móvel (MA), e ϵ_t é o erro aleatório no período t , assumido como um ruído branco com média zero e variância constante.

Neste estudo, utilizamos a série temporal de cada jogador, sendo a medida de tempo as rodadas ocorridas e os valores correspondentes à sua pontuação na respectiva rodada.

3.2. Modelo Matemático - Otimização mono-objetivo

Nesta seção, descreveremos o modelo de programação matemática utilizado para a tomada de decisão na escalação dos times no Cartola FC. O modelo emprega técnicas de otimização para selecionar os jogadores de forma a maximizar o desempenho da equipe com base nas previsões de pontuações calculadas anteriormente.

Essa abordagem permite a seleção otimizada de jogadores com base nas previsões de desempenho, considerando as restrições e objetivos definidos para a formação do time no Cartola FC.

A fim de representar o problema deste estudo como um problema de otimização mono-objetivo, são criados três conjuntos: J , P e E , onde J é o conjunto de todos os jogadores disponíveis em uma rodada específica, P é o conjunto de posições no time (atacante, goleiro, lateral, meia, técnico e zagueiro) e E é o conjunto de esquemas táticos disponíveis, sendo eles: 3-4-3, 3-5-2, 4-3-3, 4-4-2, 4-5-1, 5-3-2 e 5-4-1. Um outro conjunto suporta a ser definido é o Γ_i que representa o conjunto de todos os jogadores que podem jogar na posição $i \in P$. Uma vez tendo definido J e P , a decisão de escolher um jogador para uma determinada posição pode ser representada pela variável de decisão y_{ij} que assume valor 1 se o jogador $j \in J$ é selecionado para jogar na posição $i \in P$ e 0, caso contrário. No intuito de permitir a avaliação de diferentes esquemas táticos, criou-se a variável de decisão x_e que assume o valor 1 caso o esquema tático $e \in E$ seja selecionado e 0, caso contrário. Assim, o problema de otimização no contexto do nosso estudo pode ser modelado como:

$$\max \quad f(y) = \sum_{i \in P} \sum_{j \in J} a_j y_{ij}$$

$$\text{s.a.} \quad \sum_{i \in P} \sum_{j \in J} c_j y_{ij} \leq C, \quad (6)$$

$$\sum_{j \in \Gamma_i} y_{ij} = \sum_{e \in E} q_i^e x_e, \quad \forall i \in P \quad (7)$$

$$\sum_{i \in P} y_{ij} \leq 1, \quad \forall j \in J \quad (8)$$

$$\sum_{e \in E} x_e = 1, \quad (9)$$

$$y_{ij} \in \{0, 1\}, \quad \forall i \in P, j \in J \quad (10)$$

$$x_e \in \{0, 1\}, \quad \forall e \in E \quad (11)$$

O objetivo principal é maximizar a função $f(y)$, que representa a pontuação total da equipe. A variável a_j refere-se à pontuação do jogador j , resultante dos métodos de previsão utilizados, sujeita as seguintes restrições: a Restrição 6 reflete que a escolha de cada jogador j na variável y_{ij} está associada a um custo c_j . O somatório desses custos não pode exceder o valor disponível C , em cartoletas, para a montagem do elenco. A família de Restrições 7 indica que o somatório da variável y_{ij} deve corresponder ao número específico de jogadores q_i^e para cada posição i no time, dependendo do esquema tático x_e definido. Já a família de Restrições 8 reflete que um jogador j pode ser escolhido para no máximo uma posição i no time, garantindo que um jogador não seja selecionado para mais de uma posição. A família de Restrições 9 garante que apenas um esquema tático será escolhido. A família da Restrição 10 define o domínio da variável y_{ij} , onde y_{ij} assume o valor 1 se o jogador j é escolhido para a posição i , e 0 caso contrário. E por fim, a família da Restrição 11 define o domínio da variável x_e , onde x_e assume o valor 1 caso o esquema tático seja selecionado e 0, caso contrário.

4. Experimentos Computacionais

Os experimentos computacionais descritos neste estudo foram realizados em uma máquina com as seguintes especificações: Sistema Operacional Windows 10, com Processador (CPU) sendo AMD Ryzen 3600 e com Memória (RAM) de 32GB. Utilizando ainda o *Optimization Solver* HeXaly.

4.1. Conjunto de Dados e Pré-processamento

A análise do desempenho dos jogadores no Campeonato Brasileiro começa com a obtenção e pré-processamento dos dados do repositório GitHub "caRtola" [Gomide e Gualberto, 2022]. Este

repositório já coletou e organizou os dados, assegurando a qualidade e confiabilidade necessários para a análise. Os dados estão organizados em arquivos CSV por ano, contendo estatísticas abrangentes como gols, assistências, faltas e cartões, permitindo uma análise detalhada do desempenho dos jogadores. Essas estatísticas no âmbito do Cartola FC são cruciais para a pontuação do jogador na rodada, bem como seu custo.

Antes da análise, os dados passam por um pré-processamento para garantir consistência e qualidade, incluindo a verificação de valores ausentes, correção de erros e padronização de formatos. Jogadores marcados como “Nulo” e “Contundido” foram excluídos da análise para assegurar a integridade e precisão dos resultados.

4.2. Avaliação dos métodos de predição

Nesta seção, apresentamos a avaliação dos métodos de previsão utilizados e citados no estudo, destacando a eficácia e precisão de cada um na tarefa de estimar o desempenho futuro dos jogadores no Cartola FC.

Na aplicação de Médias Móveis, utilizamos médias dos valores passados da série temporal para suavizar flutuações e identificar tendências. Calculamos a média dos dois pontos anteriores (tamanho da janela igual a 2) para estimar o próximo ponto na série temporal, o que proporcionou os resultados mais satisfatórios resultante de um estudo empírico. O Alisamento Exponencial atribui pesos decrescentes aos valores passados, dando mais influência aos valores recentes. Utilizamos um parâmetro α igual a 0.8, também resultante empírico, para calcular a média ponderada dos valores passados, onde pesos das observações diminuem exponencialmente. Para o modelo ARMA, a seleção das ordens p e q é fundamental. Utilizamos o Critério de Informação de Akaike (AIC) [Moura, 2021] para identificar as combinações que minimizam o AIC, assegurando um balanceamento entre a complexidade do modelo e seu ajuste aos dados, evitando subajuste e sobreajuste. Os parâmetros utilizados nesta etapa do estudo podem ser observados na Tabela 1.

ARMA	$P = 2$ $Q = 3$
MMS	Window = 2
AES	$\alpha = 0.8$ $\beta = 0.2$

Tabela 1: Parâmetros utilizados

Para avaliar a performance geral dos três métodos de previsão, comparamos os valores do MSE médio obtido para todos os jogadores em todas as rodadas. Os resultados indicam que o modelo ARMA apresentou, em média, o menor MSE, seguido pelo Alisamento Exponencial e, por fim, pelas Médias Móveis. A Tabela 2 a seguir resume os resultados médios de MSE para os três métodos:

Método	MSE Médio
ARMA	3.21
Médias Móveis	8.10
Alisamento Exponencial	6.95

Tabela 2: MSE Médio dos Métodos de Previsão

Para exemplificar a utilização dos métodos, foi composto um mesmo time utilizando cada

método proposto e comparadas as pontuações previstas em cada método. Na Tabela 3 é analisado cada pontuação obtida para um time específico composto ao acaso para uma rodada.

Impacto na composição de equipes				
Jogador	MMS	AES	ARMA	Pontuação real
Tiago Volpi (gol)	5,2	5,7	7,5	8,6
Jackson (zag)	4,3	5,3	7,2	8,5
Léo Ortiz (zag)	3,2	6,2	6,8	7,5
Reinaldo (lat)	8,7	9,3	14,5	16,3
Heitor (lat)	5,8	7,3	8,3	10,3
Thiago Galhardo (mei)	7,4	8,7	9,4	9,9
Fernando Sobral (mei)	9,5	11,2	12,1	12,5
Nenê (mei)	6,8	9,6	9,5	10,6
Marinho (ata)	10,2	14,9	15,3	16,4
Keno (ata)	12,6	18,5	20,3	21,1
Abel Hernández (ata)	4,8	6,8	7,4	10,8
Jorge Sampaoli (tec)	5,6	7,2	7,1	8,1
Pontuação total	84,1	110,7	125,4	140,6

Tabela 3: Time composto baseado nos métodos de previsão

Neste exemplo, o método MMS apresentou uma pontuação total de 84,1, a menor entre os três métodos avaliados. Este método subestima significativamente as pontuações reais dos jogadores, indicando que, por ser sensível a variações pontuais, pode não capturar bem as tendências e variações complexas. Uma limitação conhecida das médias móveis simples é a equalização de peso para todas as observações dentro da janela de cálculo, o que pode levar a subestimações em caso de tendência ascendente nas pontuações. No método AES teve uma performance melhor que o MMS, com uma pontuação total de 110,7. O AES estimou as pontuações de forma mais próxima à realidade em vários casos, como no jogador Keno (ata), com uma previsão de 18,5 comparada à pontuação real de 21,1. Este método, que dá mais peso às observações mais recentes, ajusta-se melhor às mudanças na performance dos jogadores ao longo do tempo, embora ainda apresente algumas discrepâncias. Por sua vez, O método ARMA destacou-se como o mais preciso entre os três, com uma pontuação total prevista de 125,4, mais próxima da pontuação real total de 140,6. Este método combina componentes autoregressivos e de médias móveis, permitindo uma modelagem mais robusta das séries temporais das pontuações dos jogadores. A precisão do ARMA é evidente em casos como o de Reinaldo (lat), com uma previsão de 14,5 comparada à pontuação real de 16,3, e Marinho (ata) com previsão de 15,3 comparada à pontuação real de 16,4.

4.3. Resultados computacionais do *Framework*

Ao incorporar os resultados adquiridos através dos métodos de previsão a estrutura proposta do *Framework*, obtém-se assim uma composição otimizada da equipe que nos traz o maior retorno de pontuação, respeitando as Restrições e Regras do Cartola FC.

Ao comparar os resultados a pontuações totais das equipes obtidas através do *Framework* usando Médias Móveis, Alisamento Exponencial e ARMA, pode-se observar novamente a capacidade do modelo em replicar o valor bem próximo a pontuação ideal, como mostrado no gráfico da Figura 3, onde para cada rodada são observados os valores resultantes da pontuação obtida pela equipe composta através de cada método.

Para avaliar a eficácia dos possíveis componentes do nosso *Framework*, comparamos os resultados com um cenário de informação perfeita, onde as previsões correspondem exatamente aos valores reais de desempenho dos jogadores da equipe formada, servindo como *Benchmark*.

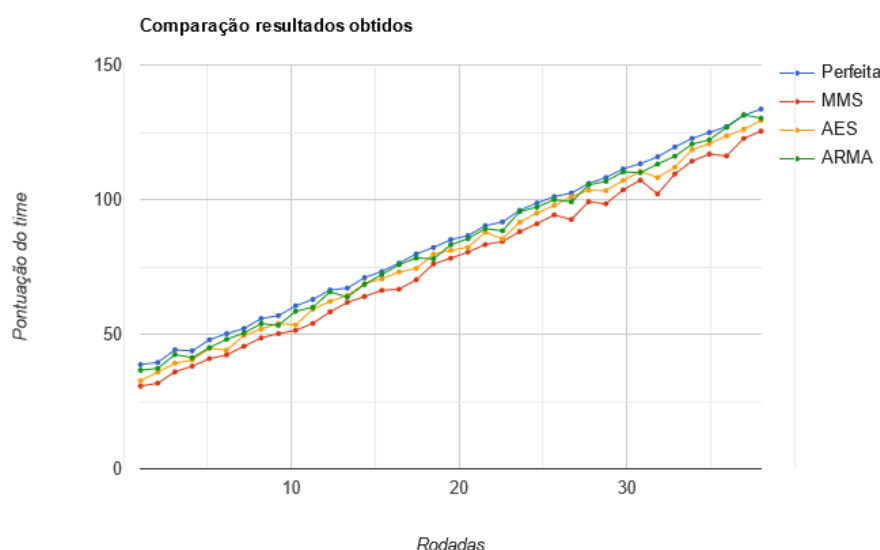


Figura 3: Comparação entre os métodos

Durante a análise, observamos que a distância máxima entre a pontuação mais alta e a mais baixa foi de 25 pontos. A maior variação ocorreu na rodada 36, onde o *Framework* usando o método ARMA alcançou uma pontuação total de 126 pontos, enquanto o *Framework* usando o método de Médias Móveis apresentou apenas 101 pontos, resultando na maior diferença entre os métodos utilizados pelo *Framework*. A menor variação ocorreu na rodada 37, onde todos os métodos utilizados pelo *Framework* tiveram um desempenho semelhante, com uma diferença de apenas 8 pontos entre eles e o resultado do *Framework* utilizando o modelo ARMA foi igual ao valor real. Além disso, nas rodadas 26 e 32, o método ARMA utilizado no *Framework* registrou aumentos significativos em suas pontuações totais, com incrementos de 10 pontos em comparação com as rodadas anteriores. Em contrapartida, o *Framework* usando o método de Médias Móveis teve quedas consecutivas nas mesmas rodadas, com diminuições de até 12 pontos. Isso destaca a capacidade do método ARMA utilizado no *Framework* de identificar e ajustar rapidamente às tendências de desempenho dos jogadores, enquanto o método de Médias Móveis falhou em capturar essas variações dinâmicas.

Nas rodadas finais, especialmente na rodada 37, o *Framework* usando o método ARMA conseguiu manter uma pontuação acumulada de 133 pontos, enquanto o método do *Framework* de Alisamento Exponencial atingiu 117 pontos e o método, também usando *Framework*, de Médias Móveis ficou com apenas 106 pontos. A maior diferença total foi observada na rodada 32, com uma discrepância de quase 25 pontos entre ARMA e Médias Móveis, consolidando a superioridade do *Framework* utilizando o método ARMA.

Mesmo não sendo o objetivo direto deste estudo, é interessante analisar como as previsões utilizadas pelo *Framework* proposto podem impactar de maneira significativa o desempenho do usuário através da variação de Cartoletas. A Figura 4 mostra o resultado da variação das cartoletas em função dos métodos usados pelo *Framework*. Quando o *Framework* utiliza o método ARMA, observa-se um maior acúmulo de Cartoletas ao longo das rodadas. Esta tendência é observada com

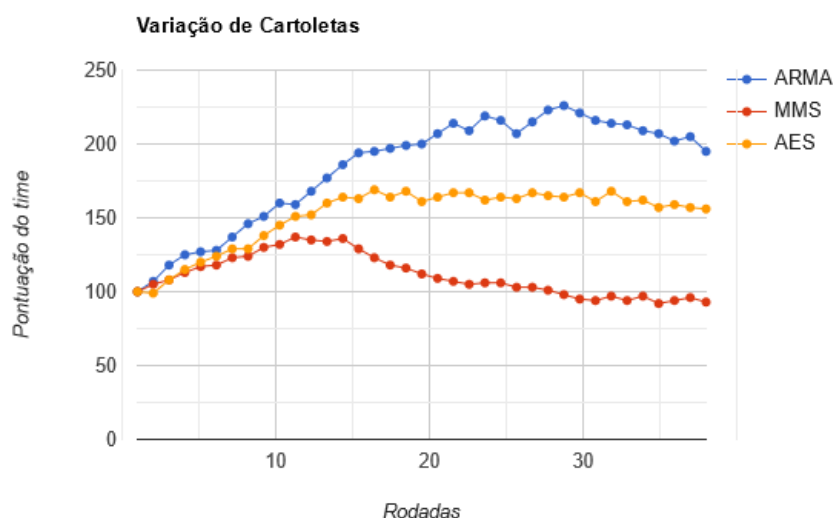


Figura 4: Variação das Cartoletas ao longo do Campeonato

maior clareza a partir da rodada 12. Isto acontece porque o método ARMA consegue prever melhor os resultados dos jogadores, implicando em uma equipe mais próxima da ideal e, consequentemente, um maior acúmulo de Cartoletas. Usando o mesmo raciocínio, como o método de Médias Móveis tende a prever a pontuação dos jogadores com maiores erros, a equipe estará mais longe da ideal e, por isso, percebe-se que a quantidade de Cartoletas acaba com um valor menor que o inicial, apesar de aumentar nas primeiras rodadas. É possível dizer que a escolha do modelo de previsão no *Framework* proposto influencia diretamente na obtenção das cartoletas pelos usuários do jogo.

5. Conclusões e Trabalhos Futuros

Os resultados obtidos neste estudo demonstram claramente a superioridade do modelo ARMA em comparação com os outros métodos avaliados. Ao analisar o desempenho dos métodos ao longo das rodadas do Cartola FC, observamos que o ARMA foi capaz de capturar com maior precisão os padrões temporais complexos presentes nos dados históricos, resultando em previsões mais assertivas.

Em termos de precisão, o modelo ARMA alcançou uma acurácia impressionante de 80%, destacando-se como a melhor opção para a previsão de desempenho dos jogadores neste contexto específico do Cartola FC. Por outro lado, o método de Médias Móveis teve o pior desempenho, com uma precisão de cerca de 40%. Isso evidencia a limitação desse método em capturar as flutuações dinâmicas do desempenho dos jogadores ao longo do tempo, o que resultou em previsões menos precisas e, consequentemente, em pontuações totais mais baixas.

Ao longo das análises realizadas, notamos que o modelo ARMA foi capaz de identificar e ajustar rapidamente às tendências de desempenho dos jogadores, como evidenciado pelos aumentos significativos em suas pontuações totais em algumas rodadas específicas. Por outro lado, os métodos de Médias Móveis e Alisamento Exponencial mostraram-se menos adaptáveis a essas variações dinâmicas, resultando em pontuações totais mais inconsistentes ao longo do campeonato. Além disso, o *Framework* desenvolvido mostrou-se eficaz ao integrar essas previsões no processo de otimização, permitindo a seleção de equipes que maximizam a pontuação esperada, respeitando as restrições de orçamento e as regras do jogo. Esse processo cíclico e contínuo de análise e atualização assegurou que as decisões fossem sempre baseadas nos dados mais recentes e precisos.

O presente estudo representa uma contribuição significativa para o campo, explorando os diferentes métodos de previsão aplicados ao contexto de *Fantasy Game*. No entanto, como toda pesquisa, ele levanta novas questões e sugere áreas promissoras para investigações futuras. Algumas propostas para a continuidade desse estudo é baseada na análise de tendências temporais e ciclos recorrentes de jogadores específicos, bem como a integração de dados contextuais externos ao Cartola FC que possam enriquecer a análise das equipes. Considerar informações como a localização geográfica dos jogos, condições meteorológicas, notícias sobre lesões, transferências de jogadores, entre outros, pode proporcionar uma compreensão mais completa do desempenho dos atletas.

Agradecimentos

Os autores agradecem a CAPES, ao CNPq e à FAPERJ pelo financiamento parcial desta pesquisa.

Referências

- Bangdiwala, M., Choudhari, R., Hegde, A., e Salunke, A. (2022). Using ml models to predict points in fantasy premier league.
- Beliën, J., Goossens, D., e Van Reeth, D. (2017). Optimization modelling for analyzing fantasy sport games. *INFOR. Information systems and operational research*, 55(4):275–294. ISSN 0315-5986.
- Budgen, D. e Brereton, P. (2006). Performing systematic literature reviews in software engineering. volume 2006, p. 1051–1052.
- Casagrande, M. C., Lavarda, S. D. L., e Silveira, A. C. M. (2021). Fantasy game e reconfiguração da cultura desportiva. *Intexto (Porto Alegre)*, (52):95133–24. ISSN 1807-8583.
- Cortinhas, D., da Silva, D. A., Mendonça, I., Amorim, G. F., e González, P. H. (2023). Otimização combinatória aplicada a composição de equipes esportivas.
- Farquhar, L. K. e Meeds, R. (2007). Types of fantasy sports users and their motivations. *Journal of computer-mediated communication*, 12(4):1208–1228. ISSN 1083-6101.
- GloboEsporte.com (2017). Globo esporte. URL <https://ge.globo.com>. Acessado 25/05.
- Gomide, H. e Gualberto, A. (2022). *CaRtola: Extração de dados da API do CartolaFC, análise exploratória dos dados e modelos preditivos em R e Python*. URL <https://github.com/henriquepgomide/caRtola>.
- Horvat, T. e Job, J. (2020). The use of machine learning in sport outcome prediction: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10:e1380.
- Moura, A. R. (2021). Seleção de modelo. URL <https://repositorio.ufpb.br/jspui/handle/123456789/23007>. Acessado 01/06.
- Robinson, C. (2020). The prediction of fantasy football. *Mathematics Senior Capstone Papers*, 20. URL <https://digitalcommons.latech.edu/mathematics-senior-capstone-papers/20>.
- Silva, M. A. (2019). Cartola fc: fantasy game, mídia e engajamento do torcedor no futebol nacional.
- Singh, N. (2020). Sport analytics: A review. *The International Technology Management Review*, 9.