

Buscando Métricas para o Impacto de um Treinador de Futebol

Marcos Vinícius de Carvalho Ribeiro, (*marcosvcr@dcc.ufmg.br*) *Ciência da Computação - UFMG,*

Abstract—This study explores the development and analysis of a metric termed “*Manager Coefficient*,” designed to quantify a football coach’s contribution to team performance. The metric was developed within the course “Data Science Applied to Football” and aimed to provide a standardized measure of a coach’s impact on match outcomes. Data was collected from FBRef and Wikipedia, encompassing Premier League matches from the 2016-2017 to 2023-2024 seasons. The study involved the formulation of two models, the *Expected Points* (xP) model, to measure how favourite a team is in a match, and another one called *Expected Points Overperformance* (xPovr). From those two models, the *Manager Coefficient* is the value the feature of who the manager was in a match had in explaining the xPovr. This value was then analyzed further as a possible metric to measure the influence a manager had in the result of a match, however, the metric’s predictive validity was found lacking, especially in predicting future trends or when applied outside the training dataset context. This study concludes that while the *Manager Coefficient* offers some insights within its specific data and training context, it does not provide reliable predictive power when generalized. Future research will explore alternative models and metrics, incorporating tactical and methodological compatibility between coaches and teams.

I. INTRODUÇÃO

Durante o estudo na disciplina de Ciência de Dados Aplicada ao Futebol, foi desenvolvida uma possível métrica para quantificar a contribuição de um treinador para os resultados de uma equipe, denominada *Manager Coefficient*. Este trabalho busca aprimorar os modelos que originam essa métrica e analisar quais informações ela realmente proporciona. Para isso, reformulou-se alguns modelos com o objetivo de criar um significado geral para a métrica e verificar se ela é válida para aplicações práticas ou se apenas se aplica ao contexto que a gerou.

A. Dados

Todos os dados das partidas foram extraídos do site FBRef, consistindo em dados de súmulas básicos. A lista de técnicos e suas respectivas equipes foi retirada da Wikipedia, para associar cada técnico a cada partida. As partidas englobam todas as disputas da Premier League da Inglaterra das temporadas de 2016-2017 até a temporada de 2023-2024, até a elaboração deste relatório.

II. METODOLOGIA

A. Análise Exploratória

Este estudo focou nos pontos adquiridos por uma equipe durante o período em que um técnico esteve no comando.

Orientador - Professor Adriano Machado, do Departamento de Ciência da Computação (DCC) da UFMG

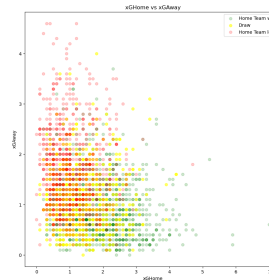


Fig. 1. Resultado (Time da casa) comparado com o xG de cada equipe

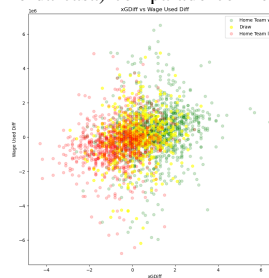


Fig. 2. xGDiff (xGHome-xGAway) e diferença na soma dos salários dos jogadores que jogaram por cada equipe

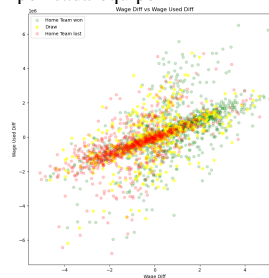


Fig. 3. Diferença na folha salarial de cada equipe e Diferença na soma dos salários dos jogadores que jogaram por cada equipe

Portanto, foi necessário analisar o espaço amostral em relação ao resultado das partidas, tentando capturar uma tendência geral. Ao observar a distribuição de resultados nas figuras 1, 2, 3, para a maioria das métricas de comparação, a separação entre vitória e derrota é clara, com empates introduzindo alguma confusão. Esses dados indicam claramente um gradiente das zonas onde o time mandante geralmente vence até as zonas onde geralmente perde. Esta clara divisão é importante, indicando que modelos simples podem prever partidas com boa precisão.

B. Expected Points

No trabalho realizado na disciplina, o modelo de Expected Points (xP) foi desenvolvido a partir de um modelo de classificação utilizando Regressão Logística Balanceada, buscando maximizar a acurácia. O xP da equipe mandante é dado pela fórmula:

$$xP = 3 * P(W) + P(D)$$

onde P(W) é a probabilidade do modelo classificar a partida como vitória do time da casa, e P(D) é a probabilidade de empate. O xP do time visitante é o complementar:

$$xP(visitante) = 3 * P(L) + P(D)$$

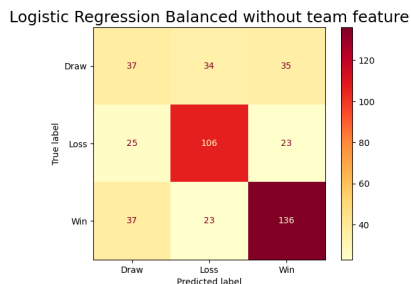
As *features* deste modelo incluem:

- xGHome: xG que o time da casa teve
- xGAway: xG que o time visitante teve
- Home Wage: Folha salarial total do time da casa
- Away Wage: Folha salarial total do time visitante
- Home Used: Soma dos salários dos jogadores que jogaram pelo time da casa
- Away Used: Soma dos salários dos jogadores que jogaram pelo time visitante
- Wage Diff: Folha salarial do mandante - Folha salarial do visitante
- Wage Used Diff: Soma dos salários dos jogadores do mandante - Soma dos salários dos jogadores do visitante
- Home Used pct: Porcentagem do salário total que foi usado pelo mandante
- Away Used pct: Porcentagem do salário total que foi usado pelo visitante

Os resultados deste modelo de classificação foram satisfatórios:

Accuracy : 0,61

F1 – score : 0,61



No entanto, o objetivo deste trabalho é minimizar o erro do xP em relação ao resultado real. Como xP é um valor numérico, tratar essa tarefa como uma tarefa de regressão é mais interessante. Por isso, o modelo foi substituído. Utilizando a função *CatBoostRegressor*, um novo modelo para o xP foi construído, utilizando as seguintes *features*:

- Home: Time mandante
- Away: Time visitante
- Home Year: Time mandante + Temporada

- Year: Temporada
- Away Year: Time visitante + Temporada
- Home Wage: Folha salarial total do time da casa
- Away Wage: Folha salarial total do time visitante
- Home Used: Soma dos salários dos jogadores que jogaram pelo time da casa
- Away Used: Soma dos salários dos jogadores que jogaram pelo time visitante
- Home Used pct: Porcentagem do salário total que foi usado pelo mandante
- Away Used pct: Porcentagem do salário total que foi usado pelo visitante

A principal diferença dessas *features* para o modelo anterior é a remoção das *features* de xG. A ideia é capturar o máximo de fatores que não envolvem a partida em si para definir o xP, sendo que o xG é uma *feature* gerada no contexto da partida jogada. Já a adição da variável controlando quem é cada time, e quem é o time e qual temporada está traz a ideia de que o modelo precisa aprender a estimar coeficientes relacionados a cada time na temporada e a cada time como um todo.

Nesse caso, dois modelos são implementados, o para o xP do time mandante e outro para o time visitante. Os resultados deste novos modelos, em termos de RMSE e R2, foram:

Home Model

RMSE : 0.65

r2 : 0.76

Away Model

RMSE : 0.78

r2 : 0.64

Este RMSE está dentro do valor aceitável, visto que toda partida termina em 3, 1 ou 0 pontos, e raramente o modelo prevê exatamente esses valores como xP. Assim como o modelo de classificação busca minimizar o número de resultados errados através da acurácia, a regressão visa minimizar a soma total do erro. Dado que a análise ocorrerá sobre períodos maiores do que uma partida isolada, o importante é minimizar o RMSE, que era o objetivo desta regressão. Pode-se afirmar que este RMSE representa, dadas as *features*, o máximo que o modelo consegue se aproximar dos resultados reais. Essa mudança de foco para minimizar o RMSE também será útil futuramente.

C. Manager Coefficient

O *Manager Coefficient* surge da tentativa de explicar a variação entre o xP e o que uma equipe realmente adquiriu em uma partida. Essa diferença entre o esperado e o real, chamada de Expected Points Overperformance (xPovr), pode ser considerada a parte que o modelo de xP não conseguiu explicar apenas com as *features* básicas. Um segundo modelo, utilizando *features* mais elaboradas, foi implementado. Esse modelo busca prever o xPovr de uma partida específica, complementando o modelo simples de xP e tentando associar

o erro do xP aos fatores utilizados como *features*. As *features* utilizadas foram:

- xGHome: xG que o time da casa teve
- xGAway: xG que o time visitante teve
- xGHome Mean: Média de xG das partidas anteriores do time da casa na temporada
- xGAway Mean: Média de xG das partidas anteriores do time visitante na temporada
- Home Manager: Treinador do time da casa
- Away Manager: Treinador do time visitante
- Home Year: Time da casa + Temporada
- Away Year: Time visitante + Temporada
- Home: Time da casa
- Away: Time visitante
- Home Last Match xPoints ovr: quanto foi o xPovr do time da casa na última partida que foi mandante
- Away Last Match xPoints ovr: quanto foi o xPovr do time visitante na última partida que foi visitante
- Home Rolling avg xPoints ovr: Média de xPovr das últimas cinco partidas do time como mandante
- Away Rolling avg xPoints ovr: Média de xPovr das últimas cinco partidas do time como visitante

Este modelo, com versão para mandante e visitante, apresentou os seguintes resultados:

Home Model

$$RMSE : 0.25$$

$$r2 : 0.67$$

Away Model

$$RMSE : 0.35$$

$$r2 : 0.67$$

Como pode ser observado, quem foi o técnico no comando da equipe na partida é uma das *features*. Utilizando SHAP para avaliar o quanto cada feature adicionou ou removeu no xPovr previsto pelo modelo, é possível extrair quanto um técnico específico contribuiu para a previsão do xPovr. Esse valor é o *Manager Coefficient* daquele treinador naquela partida. Pode-se ver nas figuras 4 e 5 para os modelos do mandante e visitante, respectivamente, um plot geral dos coeficientes dados para cada variável. Um fator importante a ser destacado é que este valor traz as correlações que cada valor desta variável categórica tem com o xPovr. Isso não implica que o técnico causou esse impacto, uma vez que o técnico também influencia o valor das outras *features*. O objetivo aqui não é analisar valores singulares, mas tentar estabelecer tendências desse valor com as outras variáveis.

D. Simulando Partidas

A ideia por trás do *Manager Coefficient* é que, ao simular uma partida hipotética de um time com um determinado técnico, deve-se encontrar qual seria o *Manager Coefficient* desse técnico nessa partida, analisando tendências e outros resultados, e em seguida incorporar esse valor no modelo de

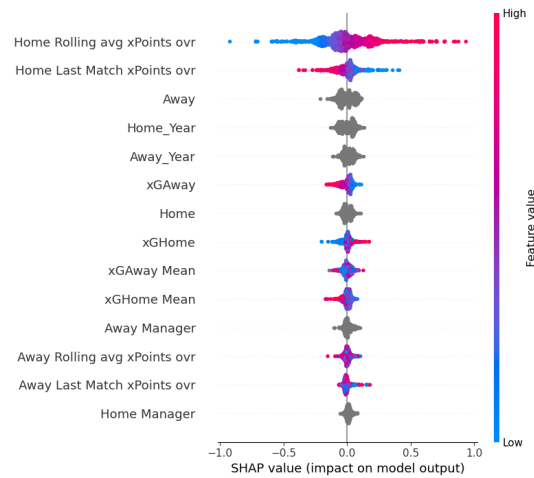


Fig. 4. Impacto de cada variável para o modelo do time mandante

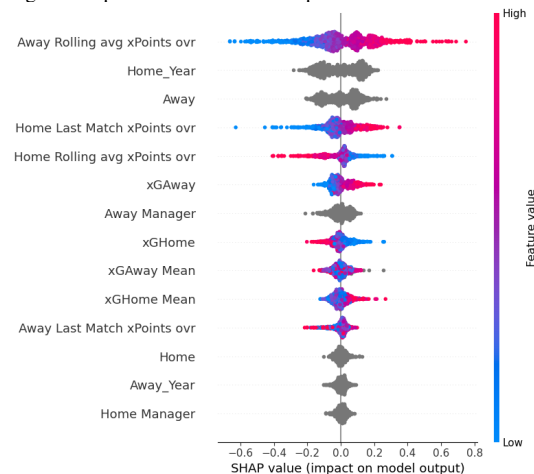


Fig. 5. Impacto de cada variável para o modelo do time visitante

previsão de resultados. Um exemplo é a métrica denominada "Change Score", que mede o quanto uma troca de técnico no meio da temporada explica a variação no desempenho da equipe. Nesse caso, valores positivos indicam trocas benéficas, enquanto valores negativos indicam trocas prejudiciais. O primeiro desafio é atribuir corretamente o valor do *Manager Coefficient*. Para o *Change Score*, inicialmente foi utilizada a média dos *Manager Coefficient* das partidas que o técnico realmente comandou no intervalo. No entanto, isso não captura bem a volatilidade dessa métrica. Neste trabalho, foi criado um modelo, utilizando uma regressão, para prever o *Manager Coefficient*. Este modelo apresentou resultados melhores do que simplesmente estimar a média. As *features* desse modelo são:

- Home Manager: Nome do técnico do time mandante
- Away Manager: Nome do técnico do time visitante
- Home: Time mandante
- Away: Time visitante
- Year: Temporada (Variável Categórica)
- Years since first year: Anos desde a primeira temporada da base de dados (Variável numérica)
- Home Manager Years in df: Anos desde a primeira vez

que o técnico do time mandante aparece na base de dados

- Away Manager Years in df: Anos desde a primeira vez que o técnico do time visitante aparece na base de dados

a separaçã das variáveis "Year" e "Years since first year" ocorreu para que o modelo consiga capturar eventuais coeficientes esporádicos locais em um ano (variável "Year"), além de um comportamento geral para a progressão dos anos. Neste trabalho, também houve uma revisão em relação ao trabalho anterior na disciplina quanto ao modelo usado para prever o resultado. Anteriormente, o modelo tentava, utilizando o *Manager Coefficient* como uma das *features*, prever o xG de cada time e, a partir disso, atribuir um novo valor de xP utilizando o modelo da Regressão Logística com o xG atualizado. A lógica por trás disso é que, como o *Manager Coefficient* afeta apenas aquela partida, faz sentido que ele influencie apenas a variável que representa a tendência local daquela partida, que é o xG. No entanto, quem é o técnico não influencia apenas isso em uma partida. Aproveitando a remoção da feature de xG no novo modelo de xP, o novo modelo para calcular o xP de uma partida fictícia é o mesmo do anterior, exceto que agora o *Manager Coefficient* é diretamente adicionado. O esperado é que, se o *Manager Coefficient* é uma métrica que contribui para explicar melhor os resultados, dois modelos similares diferindo apenas por essa feature deveriam ter o modelo com o *Manager Coefficient* apresentando melhor desempenho. Isso indicaria que essa variável está contribuindo para a avaliação final e, portanto, avaliaria se o *Manager Coefficient*, como métrica, contribui para a melhora nos resultados. Os resultados desse modelo, com a adição do *Manager Coefficient*, foram:

Home Model

$RMSE : 0.60$

$r2 : 0.79$

Away Model

$RMSE : 0.69$

$r2 : 0.72$

Comparado com o modelo inicial de xP, não houve melhora nas métricas de erro nem de R^2 . Ao analisar a *Feature Importance* desse modelo, observa-se que o *Manager Coefficient*, de forma geral, não teve importância significativa na performance do modelo, o que já é um primeiro descrédito para essa métrica.

E. Relação Temporal do Manager Coefficient

Com o *Change Score*, tem-se uma métrica para avaliar o quanto uma troca de técnico foi efetiva imediatamente. Isso, no entanto, não fornece muita informação nova, pois geralmente boas trocas são observáveis pelo desempenho na tabela. O modelo de previsão do *Manager Coefficient* consegue capturar comportamentos hipotéticos quando os dados contemporâneos a esse cenário já existem para os técnicos analisados. Os dados que o modelo treina para o técnico substituído já existem naquele ano e naquela mesma equipe, facilitando o aprendizado. Isso não é muito útil.

O grande desafio para previsões futuras é justamente entender como seria o desempenho do *Manager Coefficient* ao longo do tempo. Como seria a temporada do time X se contratasse o técnico Y no próximo ano? O que eles poderiam esperar? Para realizar tais simulações, é necessário compreender o comportamento do *Manager Coefficient* com o passar do tempo. Para isso, o modelo de previsão do coeficiente precisa ser eficaz em prever dados, principalmente em anos sem registro de partidas daquele técnico para treinar.

Infelizmente, ao realizar um teste de validação para dados futuros, o modelo apresentou um desempenho insatisfatório, com um R^2 muito baixo, indicando que ele não fornece previsões melhores do que estimar a média.

Esse teste foi realizando um split temporal entre treino e teste da seguinte maneira:

```
n ← 2017
while there are years left to process do
  Train on data where Year ≤ n
  Test R2 score on data where Year = n + 1
  n ← n + 1
end while
```

O resultado desse teste foi:

$[-0.47, 0.00, -0.38, -0.10, -0.03]$

Cada posição no vetor indica um ano n de 2017 até 2021 no algoritmo. Esse desempenho ruim indica que o modelo não consegue aprender as tendências temporais necessárias com as *features* e dados atuais. Se é possível ou não obter essa relação é uma questão para futuros trabalhos, seja com uma base de dados maior (mais ligas, mais temporadas, etc.), melhores *features* ou até modelos diferentes. O fato é que, neste trabalho, essa previsão não se mostrou eficaz. Para explicar isso, basta realizar um PCA nas *features* do modelo de xPovr. Pode-se observar que a feature do técnico, em média, não tem grande contribuição no valor previsto e, portanto, tem um comportamento bem instável e aleatório.

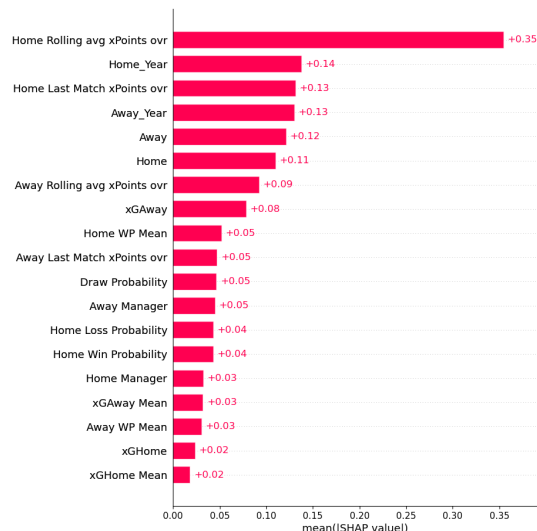


Fig. 6. *Feature Importance* para o modelo de xPovr pro time mandante

Isso permite uma grande variação entre uma execução e outra do treinamento do modelo, sem comprometer a consistência geral das previsões. Isso mostra que esse valor não possui um significado real sem os dados de treinamento do modelo original para explicá-lo. É possível ver isso ao analisar a variação que o modelo apresenta entre diferentes execuções ao prever os mesmos conjuntos de jogos. O exemplo a seguir mostra cinco diferentes execuções das partidas em casa do treinador *Jürgen Klopp*, do *Liverpool*, na temporada de 2023, além do valor que o modelo atribuiria a um técnico inexistente na base de dados.

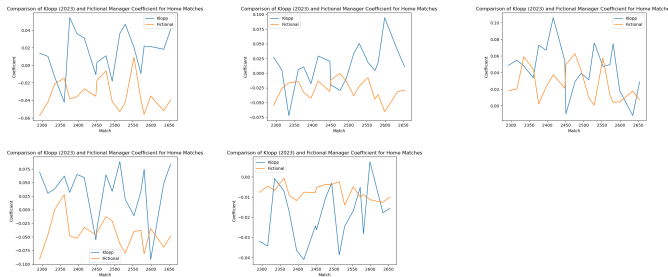


Fig. 7. Diferentes treinamentos geram valores radicalmente diferentes para o *Manager Coefficient*

Portanto, o *Manager Coefficient* pode até trazer alguma informação associada dentro do contexto do modelo de xPovr, mas qualquer tentativa de extrapolá-lo e utilizá-lo fora do contexto do modelo em que foi treinado não fornece informações úteis.

III. TRABALHOS FUTUROS

Embora o *Manager Coefficient* não tenha apresentado resultados satisfatórios, outros modelos, especialmente o de xPoints, podem ter aplicações relevantes e serão estudados mais a fundo para o desenvolvimento de novas métricas. Este estudo também ignora aspectos importantes do futebol, como as diferentes filosofias táticas e metodológicas de cada técnico. No próximo trabalho, será estudada a análise de compatibilidade entre equipe e técnico em relação a essas questões táticas, buscando determinar se isso resulta em melhores resultados.

IV. CONCLUSÃO

Foi tentado criar uma métrica inicial para analisar o quanto um treinador melhora ou piora, de forma direta e sem considerar aspectos independentes a ele, os resultados em pontos de uma equipe na tabela da Premier League. A métrica proposta, o *Manager Coefficient*, baseia-se na *Feature Importance*, extraída usando SHAP, da variável categórica que indica quem era o treinador da equipe em uma partida. Ao realizar análises mais profundas sobre esse valor, especialmente ao observar o *Feature Impact* do modelo via SHAP e os resultados do modelo de xP incluindo o *Manager Coefficient*, nota-se que ele só tem alguma utilidade dentro do contexto dos dados utilizados no treinamento. Além disso, quando tenta-se prever valores cujas *features* não estão presentes no conjunto de treino, como dados sobre anos futuros ou de técnicos que não comandaram

partidas naquele ano, o desempenho é insatisfatório, com alta variabilidade, não sendo bons indicadores. Isso mostra que a métrica, por si só, não possui correlações bem definidas individualmente, e por isso a relação temporal necessária para uma boa performance da métrica não é bem estabelecida.

REFERENCES

- [1] S. Kuper, *Soccernomics (2018 World Cup Edition) : Why England Loses ; Why Germany, Spain, and France Win ; and Why the United States— And Even China—Are Destined to Become the Kings of the World’s Most Popular Sport.* Publicaffairs, 2018.
- [2] “Football Statistics and History,” FBref.com. <https://fbref.com/>
- [3] “List of Premier League managers,” Wikipedia, Dec. 30, 2019. https://en.wikipedia.org/wiki/List_of_Premier_League_managers