

ABORDAGEM MULTICRITÉRIO PARA APOIO À DECISÃO DE JOGADORES DE FUTEBOL FIFA-20

Guilherme Lima Correa

Faculdade de Ciências Aplicadas da Unicamp – PO460A. R. Pedro Zaccaria, 1300, Limeira -
SP, 13484-350

guilimacorreia@gmail.com

Resumo

Para a disciplina de decisão multicritério o meu projeto final será uma ordenação de jogadores de futebol. A intenção é recriar o que o filme Moneyball apresenta, este se concentra nas tentativas de um diretor de criar um time competitivo, apesar da situação financeira desfavorável da equipe. Portanto, através do método/teoria de decisão MAUT e do método TOPSIS ordena-se qual é o melhor jogador a se contratar respeitando o valor da transferência e posição. Os dados foram retirados da plataforma que promove desafios de aprendizado de máquina chamada Kaggle.

Palavras-chave: MCDA, MAUT, TOPSIS, Futebol.

1. Introdução

No mundo do esporte de alto rendimento, onde profissionais valem milhões de reais, cada decisão de contratação se torna macro. No futebol o atleta Neymar possui o valor de transferência de, aproximadamente, 222 milhões de euros além de um salário proporcional [TRANFERMARKT, 2021]. Outros esportes seguem a mesma tendência, existem ligas que negociam seus atletas em valores no mesmo patamar, como por exemplo, NBA, NFL, etc.

Durante as janelas de contratações, na qual as negociações acontecem, os diretores se encontram diante da seguinte decisão: qual jogador contratar? Muitos fatores objetivos e subjetivos influenciam essa decisão, os jogadores são humanos, portanto, há fatores imprevisíveis que permeiam a essência do homem que são impossíveis de serem equacionados. Alguns jogadores podem não se adaptar ao novo clube por motivos pessoais. Porém, há fatores qualitativos que são possíveis transformar em quantitativos. Por exemplo, a habilidade de defender de um atleta pode ser baseada na quantidade de

desarmes que o mesmo faz em uma partida. Já um atacante pode ser avaliado pela quantidade de gols. Estes dados tentam demonstrar de maneira racional a qualidade do atleta.

O projeto teve como inspiração o filme MoneyBall (O homem que mudou o jogo). Neste filme um sujeito consegue mudar as políticas de contratação através de estatísticas, ou seja, os protagonistas do filme decidem qual jogador comprar baseado no que seus algoritmos apresentam independente da opinião pública.

A ideia deste projeto é responder à questão de qual jogador contratar baseado métodos de apoio a decisão (MCDA). O esporte definido para o estudo foi o futebol de campo, e as estatísticas foram retiradas da plataforma kaggle que são os dados compilados da base de dados utilizado em um jogo de vídeo game, FIFA-20. E como há um grande conjunto de jogadores e de características o estudo será feito na linguagem computacional Python. O grande desafio deste projeto, fora a decisão, é utilizar uma base de dados que não foi tratado possuindo assim diversas inconsistências.

Para ordenar quais jogadores devem ser comprados foram utilizados os métodos aditivos de agregação linear determinística chamado MAVT/MAUT com dois tipos de normalização: a 0-1 e pela soma e o TOPSIS.

2. Referência Bibliográfica

Os métodos de decisão multicritério podem ser classificados em três grupos os quais são encontrados na literatura:

1. Métodos de critério único de síntese (*agregação*)
2. Métodos de sobre classificação (*outranking*)
3. Métodos interativos

O grupo 1 e 2 são métodos discretos. Dentro do primeiro grupo encontra-se os métodos utilizados nesse projeto: MAUT e TOPSIS. Neste grupo a perspectiva do tomador de decisão de desagregada, isto é, o decisor define o sistema de preferência por completo.

A metodologia MAUT (*Multi-Attribute Utility Theory*) associa a Teoria de Utilidade em relação ao tratamento de problemas multiobjetivos, estes recebem o nome de atributos. Este método implica em obter uma função de utilidade por meio de um procedimento de

elicitação, o qual o decisor deve concordar com as condições estabelecidas. Por fim, obtém-se o valor de utilidade das consequências através de um processo de entrevista (elicitação) fundamentado nessa teoria [ALMEIDA, 2013].

O método TOPSIS (*Technique for order performance by similarity to ideal solution*) situa as alternativas em relação aos pontos de referência, ponto ideal e anti-ideal [ALMEIDA, 2013].

3. Metodologia

3.1 Tratamento do Banco de dados

A base de dados será um dataset da plataforma kaggle [KAGGLE, 2021] que são os dados compilados da base de dados utilizado em um jogo de vídeo game, FIFA-20. E como há um grande conjunto de jogadores, aproximadamente, 17125, e mais de 100 atributos plausíveis de análise o método será aplicado na linguagem computacional Python.

O banco de dados precisou ser tratado, visto que, no jogo existe situações fictícias onde jogadores já aposentados podem ser comprados. Como o foco desse projeto é recriar a decisão que os dirigentes devem tomar, esses jogadores conhecidos no jogo como lendários foram excluídos.

Como os métodos de agregação tem como premissa que os critérios sejam estatisticamente independentes foi-se necessário filtra-los, os critérios analisados são:

| | Ataque | Meio | Defesa | Pesos | |
|-----------|---------------------|---------------|-----------------|-------|-----|
| Critérios | Idade | | | 24% | Min |
| | Nota | | | 24% | Max |
| | Estatísticas Gerais | | | 20% | Max |
| | Poder de Ataque | Passes longos | Poder de defesa | 12% | Max |
| | Finalização | Resistência | Marcação | 10% | Max |
| | Acurácia | Passes Curtos | Interceptações | 10% | Max |

Diante dos critérios, apenas um deseja-se minimizar este é a idade do jogador. Entende-se que quanto mais novo o jogador é melhor para o clube investir. Os demais critérios são de maximização, ou seja, quanto maior o valor discreto do critério melhor avaliado ele está.

Outros tratamentos necessários foram necessários, visto que, o banco de dados não tratava o valor da transferência e de salário como números. Além disso, havia muitos jogadores com dados incorretos. Ao final dessa preparação, restou-se 12679 jogadores e 17 atributos. Os goleiros foram retirados do banco de dados, pois estes não apresentam características bem avaliadas, as características que definem o goleiro são mais qualitativas.

3.2 Aplicação

O Métodos de critério único de síntese (*agregação*) foram escolhidos para este problema por serem compensatórios, acredita-se que para o esporte essa metodologia faz sentido, visto que, uma habilidade de um atleta pode compensar algo que ele é defasado. O método AHP foi descartado para análises pelo fato de haver muitas alternativas e critérios, o que dificulta a aplicação. Portanto, os métodos aplicados foram: MAUT e TOPSIS.

Para seguir com o método, foi necessário confirmar a hipóteses de aplicação, como foi utilizado o banco de dados o problema envolve um conjunto finito de alternativas que são avaliadas de forma determinística em relação à um conjunto finito de critérios. Da mesma forma que a matriz de decisão é conhecida de forma determinística numa escala numérica, livre de incertezas ou riscos. Isso foi garantido com o tratamento do banco de dados. Por fim, tomou-se o cuidado para garantir que os critérios são estatisticamente independentes.

Antes de aplicar os métodos, há dois inputs necessários: valor da transferência e posição que desejasse contratar. O salário não foi levado em consideração, visto que, o banco de dados não tinha estes unificados em uma mesma moeda e eles são desatualizados, pois são informações que muitas vezes não são relevadas em sua totalidade. Sendo assim, ao definir o teto de transferência que deseja, o algoritmo seleciona apenas jogadores abaixo daquele teto para serem analisados reduzindo assim o conjunto de alternativas. Outro fator que é aplicado como uma condição de contorno é a posição. O futebol possui diversas posições para um mesmo setor, por exemplo, lateral direito e zagueiros são atletas que

compõem o sistema defensivo. Dessa forma, as posições foram divididas em apenas três setores: ataque, meio-campo, defesa.

3.2.1 MAUT

Tendo em vista que as hipóteses foram satisfeitas, a próxima etapa para a construção do método foi uma análise sobre as normalizações utilizada. A literatura ressalta a sensibilidade que o método possui em relação essa decisão do analista. Portanto, aplicou-se dois tipos de normalização:

1. Normalização 0-1

Se o critério é de Máximo:

$$X_{ij} = \frac{D_{ij} - \text{MIN}_i(D_{ij})}{\text{MAX}_i(D_{ij}) - \text{MIN}_i(D_{ij})} \quad (1)$$

Se o critério é de Mínimo:

$$X_{ij} = \frac{\text{MAX}_i(D_{ij}) - D_{ij}}{\text{MAX}_i(D_{ij}) - \text{MIN}_i(D_{ij})} \quad (2)$$

Vale ressaltar que neste tipo de normalização a melhor alternativa assume o valor 1 para o critério analisado, em quanto, a pior assume o valor 0.

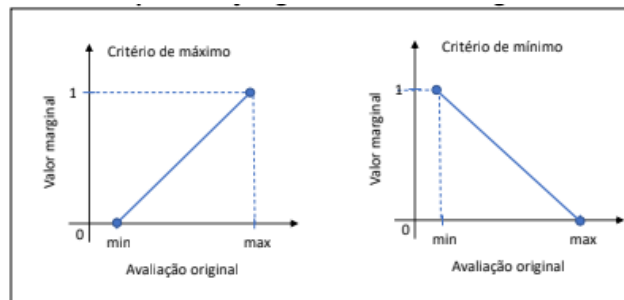


Figura 1: Representação Gráfica do Valor Marginal

2. Normalização pela Soma

Se o critério é de Máximo:

$$X_{ij} = \frac{D_{ij}}{\sum_i D_{ij}} \quad (3)$$

Se o critério é de Mínimo:

$$X_{ij} = 1 - \frac{D_{ij}}{\sum_i D_{ij}} \quad (4)$$

Feita a normalização calcula-se o Valor de Utilidade de cada alternativa com base nos pesos (W_j).

$$Vi = \sum_{j=1}^k X_{ij} * W_j \quad (5)$$

Por fim, ordena-se as alternativas de acordo com a ordenação crescente do vetor (Vi)

3.2.2 TOPSIS

Para este método foi escolhido a normalização Euclidiana.

Se o critério é de Máximo:

$$X_{ij} = \frac{D_{ij}}{\sqrt{\sum_i (D_{ij})^2}} \quad (6)$$

Se o critério é de Mínimo:

$$X_{ij} = 1 - \frac{D_{ij}}{\sqrt{\sum_i (D_{ij})^2}} \quad (7)$$

Após a normalização é necessário calcular a matriz normalizada ponderada pelos pesos **W**. Em seguida, define-se o ponto ideal e anti-ideal.

$$t_i^+ = \begin{cases} \max_i t_{ij}, & \text{se o critério é de maximizar (benefício)} \\ \min_i t_{ij}, & \text{se o critério é de minimizar (custo)} \end{cases}$$

$$t_i^- = \begin{cases} \min_i t_{ij}, & \text{se o critério é de maximizar (benefício)} \\ \max_i t_{ij}, & \text{se o critério é de minimizar (custo)} \end{cases}$$

Por fim, determina-se a distância de cada critério para cada alternativa do ponto ideal e o valor de cada alternativo é definido por:

$$Di+ = \sqrt{\sum_{j=1}^n (D_{ij+})^2} \quad (8)$$

$$Di- = \sqrt{\sum_{j=1}^n (D_{ij-})^2} \quad (9)$$

$$Ei = \frac{Di-}{(Di+) - (Di-)} \quad (10)$$

4. Resultados

Todos os métodos foram rodados alterando a faixa de preço da transferência e posição. Sendo assim, ao final do projeto tivemos 45 ordenações. As faixas de Valor de transferências definidas são:

| Faixa | Valor de transferência |
|-------|------------------------|
| 1 | 425.000 |
| 2 | 900.000 |
| 3 | 2900.000 |
| 4 | 542.000.000 |
| 5 | 1.055.000.000 |

Tabela 1: Valores de transferências disponíveis

Para avaliar os resultados, escolheu-se a última faixa de valor de transferência, pois esta possui os jogadores mais populares.

Para a posição de defensores, a ordenação dos três métodos avaliados foram:

| MAUT (0-1) | MAUT (SOMA) | TOPSIS |
|---------------------|---------------------|---------------------|
| T. Alexander-Arnold | V. van Dijk | M. de Ligt |
| V. van Dijk | M. de Ligt | T. Alexander-Arnold |
| M. de Ligt | T. Alexander-Arnold | A. Wan-Bissaka |
| A. Laporte | A. Laporte | Éder Militão |
| A. Robertson | A. Wan-Bissaka | Everticinho |
| A. Wan-Bissaka | C. Lenglet | C. Lenglet |
| Ricardo Pereira | M. Škriniar | Ronaldo Esler |
| C. Lenglet | Marquinhos | J. Gomez |
| D. Alaba | R. Varane | L. Hernández |
| Grimaldo | A. Robertson | Renan Lodi |

Figura 2: Comparação dos defensores com o maior valor de transferência

Pode-se ver que apenas 2 jogadores possuem a mesma posição em 2 métodos. O primeiro foi o *A. Laporte*, ambos nas normalizações do MAUT. E o *C. Lenglet*, ocupa a sexta posição no MAUT (normalização pela soma) e TOPSIS.

Outro ponto que vale destacar, são os jogadores que não se repetem em nenhum dos métodos, ao total foram: 12.

Para a posição de meias, têm-se:

| MAUT (0-1) | MAUT (SOMA) | TOPSIS |
|---------------------|---------------------|------------------|
| Bruno Fernandes | Bruno Fernandes | F. Valverde |
| J. Kimmich | J. Kimmich | F. de Jong |
| K. De Bruyne | F. de Jong | Rosberto Dourado |
| F. de Jong | K. De Bruyne | K. Havertz |
| F. Valverde | F. Valverde | Cucurella |
| Casemiro | Rosberto Dourado | A. Hakimi |
| S. Milinković-Savić | D. van de Beek | D. van de Beek |
| Fabinho | S. Milinković-Savić | Bruno Fernandes |
| P. Pogba | P. Pogba | J. Kimmich |
| D. van de Beek | K. Havertz | E. Camavinga |

Figura 3: Comparação dos meias com o maior valor de transferência

O MAUT, independente da normalização, convergiu em 4 posições, tanto o primeiro colocado como o segundo estão incluídos. Já o TOPSIS apresentou 7 jogadores que aparecem em outro método, mas todos em posições diferentes. Neste setor do campo analisado, 6 jogadores não apareceram em pelo menos dois métodos.

Por fim, no setor de ataque os jogadores que receberam destaque são:

| MAUT (0-1) | MAUT (SOMA) | TOPSIS |
|-------------------|-------------------|------------------|
| K. Mbappé | K. Mbappé | K. Mbappé |
| H. Kane | H. Kane | L. Martínez |
| S. Mané | L. Martínez | E. Haaland |
| R. Lewandowski | S. Mané | João Félix |
| L. Martínez | R. Lewandowski | Matheus Bardeira |
| A. Griezmann | A. Griezmann | H. Kane |
| Cristiano Ronaldo | Cristiano Ronaldo | Gabriel Jesus |
| L. Messi | T. Werner | T. Werner |
| E. Haaland | Matheus Bardeira | Richarlison |
| Neymar Jr | E. Haaland | V. Osimhen |

Figura 4: Comparação dos atacantes com o maior valor de transferência

Este foi o único setor do campo que um jogador (*K. Mbappé*) ocupou a mesma posição nos três cenários analisados, além dele, há jogadores que ocuparam a mesma posição em 4 cenários (*T. Werner* , *Cristiano Ronaldo*, *A. Griezmann* , *H. Kane*). Porém, há 6 jogadores que não se repetem em nenhum cenário.

Essas diferenças ocorrem, pois, a alteração do tipo transformação de escalas, utilizadas para a normalização, gera uma distorção dos dados quando desvinculada da escolha das

constantes de escala. Ou seja, ao se determinar a escolha dos dados deve buscar novos pesos para os critérios, assim é possível conseguir a mesma ordenação em todos os cenários. Portanto, nota-se que a sensibilidade dos Métodos de critério único de síntese (*agregação*) em relação a normalização é elevada e isto pode vir a ser um problema para escalonar o projeto, caso apresente para um diretor o ideal seria optar pelos mesmo jogadores em todos cenários ao qual o salário e a posição são definidos.

Vale destacar que a definição dos pesos de cada critério foi feita pelo decisor, mas esta não deve levar apenas a importância do critério independentemente. Deve-se avaliar o critério e considerar as consequências que o peso atribuído a ele pode acarretar. Caso as alternativas apresente valores muito discrepantes para um critério a normalização 0-1 pode acentuar essa diferença.

5. Conclusão

O objetivo principal do projeto era avaliar os métodos MAUT (normalização: 0-1 e soma) e o TOPSIS no ordenamento de jogadores tendo como condições de contorno o valor da transferência e a posição. Pode-se dizer que o resultado obtido apesar de possuir ordenações plausíveis para os diferentes cenários, não seria a ideal para uma situação de real de mercado, pois a decisão não teria apoio de mais de um modelo.

Portando, a escolha da normalização deve ser cuidadosa, visto que, os Métodos de critério único de síntese (*agregação*) são bem sensíveis nesse quesito.

Uma alternativa para complementar o projeto seria adicionar a avaliação de um método não compensatório, como o DEA, por exemplo. Apesar de termos muitas alternativas para a decisão, esse método não avalia pares de alternativa, ou seja, não seria inviável a aplicação dele nesta análise.

6. Bibliografia

ALMEIDA, Adiel Teixeira de. Processo de decisão nas organizações: construindo modelos de decisão multicritério. São Paulo. Atlas, 2013

KAGGLE. Data. Disponível em: <[https://www.kaggle.com/ekrembayar/fifa-21-complete-player-banco de dados](https://www.kaggle.com/ekrembayar/fifa-21-complete-player-banco-de-dados)> Acesso em: 05 de jan. de 2021.

TRANSFERMARKT. Valor de mercado. Disponível em: <<https://www.transfermarkt.com.br/neymar/marktwertverlauf/spieler/68290>> Acesso em: 05 de jan. de 2021.