

UM MÉTODO DE AJUSTE DE PESOS EM MÉTODOS DE APOIO À DECISÃO MULTICRITÉRIO BASEADO EM MEDIDAS DE EQUIDADE

Mércia de Sousa Rocha

FT-Unicamp - Rua Paschoal Marmo, 1888 - Jardim Nova Italia, Limeira - SP

mercia.s.rocha@gmail.com

Pedro Cardoso Ishara

FCA-Unicamp - Rua Pedro Zaccaria, 1300, Limeira - SP

p242947@dac.unicamp.br

Guilherme Dean Pelegrina

Universidade Presbiteriana Mackenzie - Rua da Consolação, 930 - Higienópolis - São Paulo - SP

guipelegrina@gmail.com

Leonardo Tomazeli Duarte

FCA-Unicamp - Rua Pedro Zaccaria, 1300, Limeira - SP

leonardo.duarte@fca.unicamp.br

RESUMO

A problemática da presença de vieses em conjuntos de dados tem chamado a atenção e ganhado espaço nas discussões científicas relacionadas à análise de dados. No contexto de ordenamento de dados utilizando Métodos de Apoio à Decisão Multicritério, a busca por ordenamentos ditos imparciais, capazes de mitigar o efeito de vieses frente a grupos específicos de alternativas, pode ser feita por meio do ajuste de pesos ligados aos critérios considerados em um determinado problema, uma vez que os pesos influenciam a ordem das alternativas que influencia a equidade resultante do ordenamento. Sendo assim, neste estudo, apresentamos um novo método de ajuste de pesos com base na otimização da medida de equidade conhecida como NDKL (*Normalized Discounted Cumulative KL-divergence*). Nosso método, testado em dados sintéticos e em dados reais, consiste na definição de pesos w que possibilitem um ordenamento equitativo à partir da otimização de NDKL em função desses pesos.

PALAVRAS CHAVE. Equidade. Ajuste de pesos. Métodos de Apoio à Decisão Multicritério.

ABSTRACT

In data analysis, the presence of bias in datasets have an important space in scientific discussions nowadays. In the ranking context, using Multi-Criteria Decision Analysis, the search of fair rankings, capable of mitigating biases effect involving specific groups of alternatives, can be done by adjusting the criteria weights considered in a certain problem, since the weights influence the order of the alternatives which influences the resulting ranking fairness measure. In this study, we present a new weight adjustment method based on the optimization of the fairness measure known as NDKL (*Normalized Discounted Cumulative KL-divergence*). Our method, tested on synthetic and real data, consists of defining weights w that enable fair rankings based on the NDKL optimization in function of these weights.

KEYWORDS. Fairness. Weight adjustment. Multi-Criteria Decision Analysis.

1. Introdução

O avanço tecnológico vem impulsionando e viabilizando a evolução de diversas áreas que compõem o macro ambiente da vida humana. No campo da Inteligência Artificial (IA), a partir de computadores cada dia mais potentes, é possível realizar rápidas análises de grandes massas de dados, importantes aliadas em processos decisórios. Desde o meio corporativo até as áreas médicas, como discutido em [Nazemi e Fabozzi, 2024] e em [Dharmarathne et al., 2024], é comum encontrar processos decisórios que passam pelo âmbito do aprendizado de máquina, buscando, com isso, mitigar os erros humanos envolvidos na tomada de decisão para, assim, alcançar maiores eficiência e eficácia no meio em que são aplicados.

No entanto, a presença de vieses em algoritmos de IA tem chamado a atenção por suscitar questões importantes no campo da ética e, em particular, trazendo noções de equidade como um importante assunto científico nesta área. Entre os vieses, podemos citar como exemplos [Mehrabi et al., 2021] a falta de representatividade de um determinado grupo em um conjunto de dados, a presença de erros de medida ou ainda a presença de vieses históricos, como viés de raça ou de gênero. De modo geral, o problema evidencia a necessidade de evitar que a discriminação e o preconceito inerentes ao mundo real e, portanto, enraizados nos dados sejam exportados aos modelos criados a partir deles e reproduzam os problemas sociais existentes, ou seja, continuem gerando preconceito e perpetuando a marginalização de determinados grupos, como negros e mulheres, denominados, neste estudo, como grupo sensível.

Nos campos da classificação, regressão e recomendação, ainda que não haja um consenso na área [Mehrabi et al., 2021; Sang Ah Lee e Singh, 2021], são encontradas propostas que abrangem ângulos diversos do problema, como medidas de equidade, algoritmos de auditoria para modelos de previsão, metodologias de otimização multi-objetivo [Sang Ah Lee e Singh, 2021], ou ainda, discussões sobre interpretabilidade e transparência [Carvalho et al., 2019].

Para além das áreas supracitadas, mas ainda no campo decisório, podendo envolver igualmente a análise de um conjunto de dados, temos também os Métodos de Apoio à Decisão Multi-critério (MCDA, do inglês *Multi-Criteria Decision Analysis*), amplamente utilizados na comparação de alternativas considerando critérios que podem ser conflitantes [Alvarez et al., 2021]. Sabendo que, frequentemente, essas alternativas tratam-se de pessoas ou estão direta ou indiretamente relacionadas à elas, como instituições, regiões geográficas ou países, por exemplo, as questões de equidade são cabíveis e devem ser cuidadosamente estudadas neste âmbito.

Diante desse panorama, no presente trabalho, abordamos a problemática de equidade no contexto de ordenamento de alternativas, tarefa esta central nas abordagens de MCDA. Alguns trabalhos apresentados na comunidade de computação abordam a questão de ordenamentos imparciais a partir da proposição de métricas de equidade para ordenamentos [Yang e Stoyanovich, 2017] ou soluções que buscam equidade considerando um conceito conhecido como grupo top- k [Zehlike et al., 2017]. Todavia, há ainda uma lacuna importante que diz respeito a como abordar a busca por ordenamentos ditos imparciais no contexto de MCDA.

Um elemento central nesse contexto é a etapa de agregação, presente em diversas abordagens de MCDA. De fato, nesta etapa, há certos parâmetros de modelo de agregação, que serão denominados simplesmente “pesos” (note que este termo em MCDA pode suscitar interpretações equivocadas; ver [Roy, 1996] para uma discussão aprofundada). Tais pesos influenciam a posição ocupada por cada alternativa no ordenamento e, portanto, no conjunto de alternativas que ocupam as primeiras posições e que são selecionados para um determinado fim, que pode ser o recebimento de algum tipo de benefício. Diante disso, no presente trabalho, propomos um novo método de ajuste de pesos de um modelo de agregação em MCDA com vistas à obtenção de um ordenamento que seja o mais imparcial possível, ou seja, para o qual seja mitigado o efeito do viés, quantificado por

uma métrica dita de equidade, entre dois sub-grupos de alternativas. Em nossa proposta, consideramos a otimização da medida de equidade NDKL (do inglês, *Normalized Discounted Cumulative KL-divergence*) por meio de uma meta-heurística conhecida como Evolução Diferencial.

Com relação à organização do trabalho, na Seção 2 abordaremos os principais conceitos de MCDA utilizados neste trabalho, bem como a medida de equidade e o método de otimização escolhidos. Na Seção 3, apresentaremos nossa proposta para ajuste de pesos levando em conta a otimização da métrica de equidade escolhida. Na Seção 4, discutiremos como nossa ideia é capaz de reduzir os efeitos de vieses presentes em um conjunto de dados gerado sinteticamente e também discutiremos os resultados obtidos com os experimentos numéricos em dados reais e, finalmente, na Seção 5, apresentaremos as conclusões concebidas a partir da análise dos resultados e as perspectivas de continuidade deste trabalho motivadas pelas mesmas.

2. Contexto e elementos teóricos da proposta

Nesta seção, apresentaremos os principais conceitos, essenciais para o desenvolvimento e, portanto, para a compreensão deste estudo. Entre os elementos que compõem nossa proposta, começaremos por um breve panorama de MCDA, seguido de uma discussão sobre a medida adotada para avaliação de um ordenamento em termos de equidade. E finalmente, apresentaremos o algoritmo de otimização escolhido com uma visão geral do fluxo considerado em nossa proposta.

2.1. Apoio à Decisão Multicritério

MCDA refere-se a um termo genérico que engloba diferentes paradigmas de apoio à tomada decisão, de acordo com as preferências de um ou mais decisores, comparando alternativas com critérios que podem ser conflitantes [Alvarez et al., 2021]. Isso pode acontecer devido ao fato de tais critérios, tidos com um fragmento da realidade, envolverem informações de origens, interações e influências diversas, que devem ser levadas em consideração como, por exemplo, finanças, recursos humanos, aspectos éticos ou ambientais, entre outros, com os quais o decisor se depara, frequentemente em um processo decisório [Roy, 2005]. Uma vez que o conflito pode impor barreiras a uma tomada de decisão efetiva, o apoio à decisão consiste na busca, por meio de modelos explícitos, de evidências e esclarecimentos consistentes para estruturação de uma determinada solução recomendada, conferindo-lhe coerência em torno do objetivo estabelecido [Roy, 2005].

Neste estudo, entre as diferentes perspectivas abrangidas em MCDA, direcionamos a atenção a um método de agregação linear por meio do qual as alternativas são ordenadas da melhor para a pior por meio da média ponderada (SAW, do inglês *Simple Additive Weighting*). Considerando $C = \{c_1, c_2, \dots, c_m\}$ como o conjunto de m critérios e $A = \{a_1, a_2, \dots, a_n\}$ como o conjunto de n alternativas, e sendo que o vetor $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{im}]$ contém as avaliações para a i -ésima alternativa considerando todos os critérios, o valor global de cada alternativa é calculado da seguinte maneira:

$$u_i = \sum_{j=1}^m x_{ij}w_j, \quad (1)$$

onde $\mathbf{w} = [w_1, w_2, \dots, w_m]$ é o valor que representa os pesos para cada critério, tal que $w_j \geq 0$ para todo j e $\sum_{j=1}^m w_j = 1$. Dessa forma, considerando a presença de vieses, em nosso estudo o conjunto de alternativas foi dividido dois grupos: o Grupo Sensível (GS) representando o grupo pertencente a uma determinada classe afetada por vieses e o Grupo Não Sensível (GNS) representando o grupo não afetado por vieses ou até mesmo privilegiado.

Neste caso, é importante observar que a elicitação de preferências em relação aos pesos $\mathbf{w}$ trata-se da definição de pesos que cada critério terá no cálculo do valor de cada alternativa, de acordo com a preferência e objetivo do decisor. Para isso, pode ser usado um método específico, como o AHP (do inglês *Analytic Hierarchy Process*) [Saaty, 1990] ou, uma escolha mais subjetiva, levando

em consideração as predileções das partes envolvidas. Finalmente, as alternativas são ordenadas de acordo com a ordenação decrescente das utilidades, $\mathbf{u} = [u_1, u_2, \dots, u_n]$, obtidas a partir de seus valores individuais e dos pesos de cada critério. Uma vez que os valores individuais na vida real não podem ser alterados de maneira direta e instantânea, resta apenas alterar os valores dos pesos, neste caso, na tentativa de incluir a equidade como um objetivo e alterar o ordenamento nesta direção, sendo esta a abordagem que consideramos no direcionamento de nosso método.

2.2. Medida de Equidade

O conceito de equidade admite diferentes vertentes e interpretações, partindo de diferentes visões sobre igualdade e justiça entre indivíduos ou igualdade de oportunidade, como discutido, por exemplo, em [Sandel, 2011]. A própria definição de equidade no domínio da análise de dados, sabendo que foram propostas mais de 20 definições para o termo nos últimos anos [Verma e Rubin, 2018], já se impõe como desafio na área. Nesse meio existem diferentes medidas de equidade, que atendem uma extensa gama de propósitos e utilizam técnicas diversas como rKL (*Normalized Discounted KL-Divergence*) [Yang e Stoyanovich, 2017], Skew [Cem Geyik et al., 2019], EGR (*Expected Gain Ratio*) [Tahery et al., 1996], entre outras. Por sua vez, a medida NDKL (*Normalized Discounted Cumulative KL-Divergence*) [Cem Geyik et al., 2019], utilizada neste estudo, é uma medida de equidade de grupo, cujos autores defendem que um ordenamento equitativo (ou imparcial) é aquele cuja distribuição de candidatos obedece a mesma distribuição de sua população original. Esta medida oferece a vantagem de penalizar ordenamentos que são menos equitativos nas primeiras posições do ordenamento que está sendo avaliado, além de ser uma medida de fácil interpretação. Ela é definida como:

$$NDKL(\tau_r) = \frac{1}{Z} \sum_{k=1}^{|\tau_r|} \frac{1}{\log_2(k+1)} d_{KL}(D_{\tau_r^k} || D_r), \quad (2)$$

onde τ_r é uma lista ordenada de alternativas, τ_r^k são os primeiros k itens desta lista, Z é dado por:

$$Z = \sum_{k=1}^{|\tau_r|} \frac{1}{\log_2(k+1)}, \quad (3)$$

e a medida $d_{KL}(D_1 || D_2)$ refere-se à Divergência de Kullback-Leibler [Kullback e Leibler, 1951] entre as distribuições D_2 e D_1 para cada grupo g presente na população de alternativas, dada por:

$$d_{KL}(D_1 || D_2) = \sum_g D_1(g) \ln \frac{D_1(g)}{D_2(g)}. \quad (4)$$

A divergência de Kullback-Leibler mede a divergência entre duas distribuições de probabilidade de forma não simétrica, podendo ser interpretada da seguinte maneira: o valor desta divergência tende a 0 quanto mais próximas forem essas duas distribuições; por outro lado, o valor dessa divergência aumenta quanto mais diferentes forem as duas distribuições de probabilidade analisadas. Neste sentido, a função NDKL se comporta de maneira análoga, seu valor resultante estará próximo de 0 em um ordenamento dito equitativo, apenas quando as distribuições de probabilidade dos diferentes grupos existentes na população total de alternativas estiverem proporcionalmente representadas entre aquelas selecionadas em um determinado conjunto das k primeiras alternativas.

2.3. Evolução Diferencial

A Evolução Diferencial (DE, do inglês *Differential Evolution*), proposta por Price e Storm [Price e Storm, 1995], é uma ferramenta de otimização de uma função objetivo, considerando restrições, e por meio de um conjunto de heurísticas (a DE pode ser considerada uma meta-heurística). O processo de otimização parte de uma população de soluções iniciais aleatórias, que

funcionam como um mapeamento da região de busca. Nesse processo, são geradas novas soluções a partir de soluções existentes com o objetivo de direcionar a solução a um ponto de mínimo ou de máximo, dependendo do problema, aumentando as chances de encontrar um ponto ótimo global.

O processo de otimização, utilizando o algoritmo DE, consiste nas etapas de inicialização, mutação, cruzamento e seleção [Price et al., 2005]. Uma vez definidos os parâmetros do algoritmo, como número de iterações, critério de parada, taxas de mutação e de cruzamento, são estabelecidas as restrições, como limites inferiores ou superiores, e uma população de soluções aleatórias é gerada, encerrando a etapa de inicialização. Em seguida, as etapas de mutação e cruzamento correspondem à perturbação e combinação entre si de pares de soluções iniciais ou de gerações anteriores. As melhores soluções são conservadas, durante a etapa chamada de seleção, e utilizadas novamente para gerar novas soluções por mutação e cruzamento, formando novas gerações de soluções no processo evolucionário. O processo se repete até que seja atingida a quantidade de iterações desejada ou algum outro critério de parada.

Desde a sua criação, o algoritmo DE, em comparação com outras meta-heurísticas, mostrou-se versátil e efetiva na resolução de problemas de otimização de diferentes áreas, podendo envolver funções complexas com consideráveis quantidades de variáveis [Price et al., 2005]. Além disso, sua facilidade de interpretação facilita no processo de resolução de um problema.

3. Proposta: ajuste de pesos baseado na otimização de NDKL usando o algoritmo DE

Nossa proposta para definição de pesos visando um ordenamento imparcial se baseia minimização da métrica NDKL, ou seja, NDKL torna-se a função objetivo que deve ser minimizada em função de w e, para isso, utilizamos a meta-heurística DE com o intuito de viabilizar a otimização de uma função de custo de alta complexidade, uma vez que pode apresentar descontinuidades decorrentes das alterações de ordenamentos. Matematicamente, nossa proposta pode ser formulada como o seguinte problema de otimização:

$$\begin{aligned}
 \min_{\mathbf{w}} \quad & \frac{1}{Z} \sum_{k=1}^{|\tau_r|} \frac{1}{\log_2(k+1)} d_{KL}(D_{\tau_r^k} || D_r) \\
 \text{s.t.} \quad & w_j \geq 0 \quad \forall j; \\
 & \sum_{j=1}^m w_j = 1.
 \end{aligned} \tag{5}$$

Como os pesos influenciam a medida de equidade apenas de forma indireta, o algoritmo DE, aplicado a este problema, gera uma população inicial aleatória de vetores de pesos de tamanhos referentes à quantidade de critérios. Após esta etapa de inicialização, esses pesos são alterados, par a par, e combinados entre si, nas etapas de mutação e cruzamento. E, por último, a equidade é medida em função dos novos vetores de pesos e seus ordenamentos resultantes, sendo que aqueles vetores relacionados aos menores valores resultantes de NDKL são armazenados para gerar as próximas soluções, ou seja, a próxima geração de vetores de pesos no processo evolutivo da meta-heurística. Ao final do processo de otimização, o algoritmo apresentará como resultado final o vetor de pesos cujo NDKL resultante foi o menor entre todos os vetores candidatos de pesos gerados durante todo o processo, ou seja, aquele cujo ordenamento resultante é composto por alternativas cuja distribuição mais se aproxima da distribuição populacional para todos os grupos considerados no problema.

Partindo do princípio de que diversas combinações de pesos podem resultar em um valor satisfatório ou ótimo da medida de equidade, é importante que a otimização de NDKL seja repetida diversas vezes para posterior cálculo das médias e desvios-padrões dos pesos resultantes de cada otimização. Em seguida, podemos inspecionar o box-plot para verificar a variabilidade dos pesos obtidos e direcionar a etapa seguinte, que consiste no ajuste de w obedecendo essa informação e as médias de pesos obtidas. Por último, é importante calcular a equidade resultante do ajuste de

pesos com intuito de verificar se a escolha foi efetiva. A Figura 1 apresenta as etapas que resumem o método proposto.

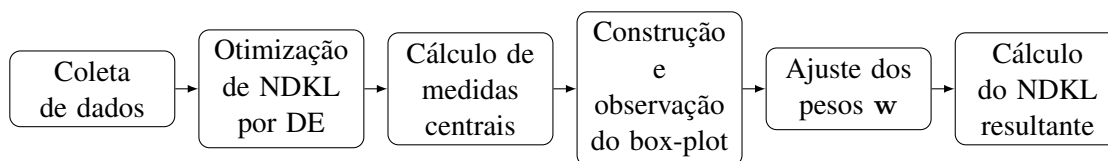


Figura 1: Fluxograma do processo de ajuste de pesos por meio da otimização de NDKL.

4. Experimentos numéricos

Para validação de nossa proposta, a realização dos experimentos numéricos foi estruturada em duas etapas: uma primeira etapa usando dados sintéticos e a segunda etapa usando dados reais. Apresentaremos primeiramente o processo utilizado para geração de dados sintéticos e os resultados obtidos com os experimentos a partir desses dados. Em seguida, apresentaremos os resultados e discussões sobre os experimentos com dados reais.

4.1. Primeira etapa: avaliação da otimização de NDKL com dados sintéticos

Nesta etapa, consideramos a simulação de diferentes conjuntos de dados onde cada conjunto foi gerado a partir de distribuições normais truncadas com valores variando no intervalo $[0, 10]$, com média, $\mu = 5$, e desvio padrão, $\sigma = 1$, com exceção de critérios enviesados, cujas médias sofreram variações tanto para as alternativas do grupo sensível quanto para aquelas do grupo não sensível. Além das médias, variamos igualmente a quantidade de critérios, sendo $m = 3$ e $m = 7$, alterando a quantidade de critérios enviesados, sendo 1 e 2 critérios enviesados para cada caso.

No caso dos critérios considerados como enviesados, foram feitas diferentes simulações, alterando as médias das populações de GS e GNS e variando, assim, o intervalo entre elas. Para o grupo não sensível, as médias estabelecidas foram $\mu = \{5, 6, 7\}$. Em contrapartida, para o grupo sensível as médias estabelecidas foram $\mu = \{3, 4, 5\}$. A variação dessas médias foi feita individualmente, somando, no caso GNS, ou subtraindo, no caso de GS, uma unidade a partir da média $\mu = 5$. Dessa forma, os intervalos resultantes entre as médias dos grupos foram de $int = \{0, 1, 2, 3, 4\}$, sendo 0 para ambos os grupos com médias 5, sem nenhum efeito de vieses e 4 para grupos em realidades opostas, sendo GS com média 3 e GNS com média 7.

A seguir, apresentaremos os resultados de otimizações consecutivas, sendo 30 repetições do processo de otimização pelo algoritmo DE, para 1 e 2 critérios enviesados, para cada intervalo entre médias de cada conjunto de critérios proveniente de conjuntos de dados com um total de $n = 2000$ alternativas, igualmente distribuídas entre os grupos, ou seja, com uma distribuição de probabilidade de 50%.

4.1.1. Otimização de NDKL para 1 critério enviesado

As Figuras 2 e 3 apresentam os box-plots dos pesos obtidos ao final de 30 repetições para cada intervalo entre médias estabelecido para o critério enviesado, neste caso w_3 para 3 critérios e w_7 para 7 critérios. Já as Figuras 4 e 5 apresentam os box-plots de NDKL resultantes para os mesmos processos de otimização. Nessas figuras, podemos observar que o peso associado ao critério enviesado, w_3 ou w_7 , tende a 0 à medida que aumenta o intervalo entre as médias de GS e GNS para o critério enviesado. Além disso, as médias de NDKL, igualmente próximas de 0, indicam que o processo de otimização foi efetivo.

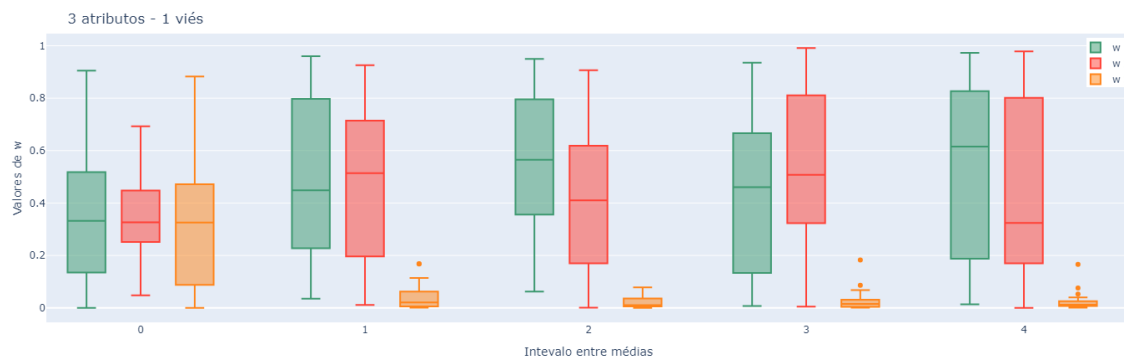


Figura 2: Box-plots dos pesos para cada intervalo entre médias no caso de 3 critérios e 1 critério enviesado.

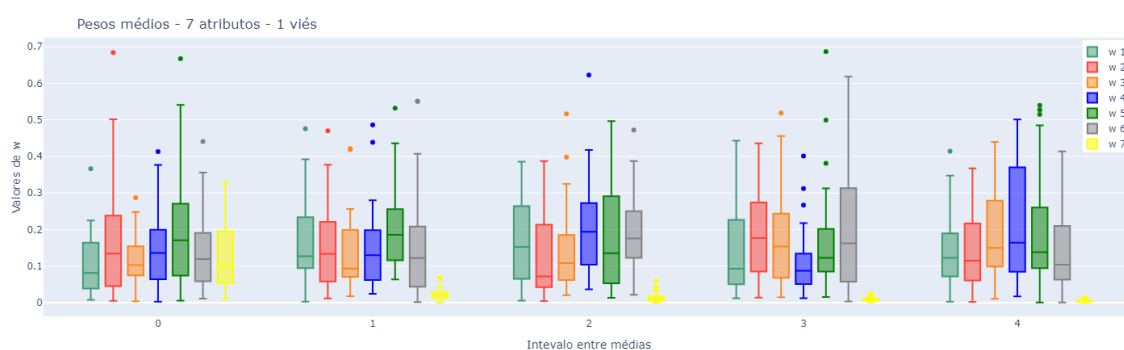


Figura 3: Box-plots dos pesos para cada intervalo entre médias no caso de 7 critérios e 1 critério enviesado.

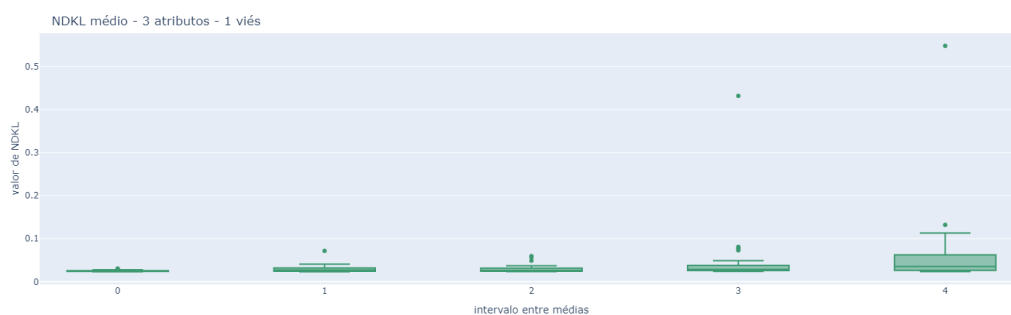


Figura 4: Box-plots de NDKL para cada intervalo entre médias no caso de 3 critérios e 1 critério enviesado.

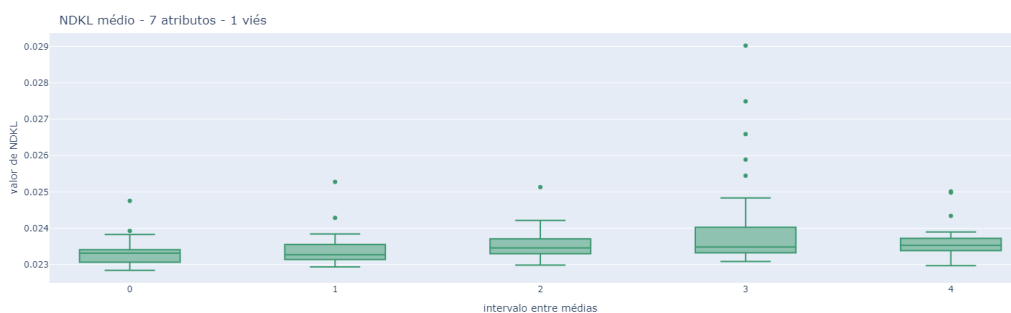


Figura 5: Box-plots de NDKL para cada intervalo entre médias no caso de 7 critérios e 1 critério enviesado.

4.1.2. Otimização de NDKL para 2 critérios enviesados

Repetindo o mesmo cenário de testes para 1 critério enviesado, as Figuras 6 e 7 apresentam os box-plots dos pesos obtidos ao final de 30 repetições para cada intervalo entre médias estabelecido para o critério enviesado. Nestes casos, bem como nos resultados obtidos com 1 critério enviesado, podemos observar que os pesos associados aos critérios enviesados, w_2 , w_3 , w_6 e w_7 , tendem a 0 à medida que aumenta o intervalo entre as médias de GS e GNS para ambos os critérios enviesados. Já as Figuras 8 e 9 apresentam os box-plots de NDKL resultantes para os mesmos processos de otimização, conservando um cenário onde essas médias estão próximas de 0, tanto para 3 quanto para 7 critérios, indicando que o processo de otimização também foi efetivo nestes casos.

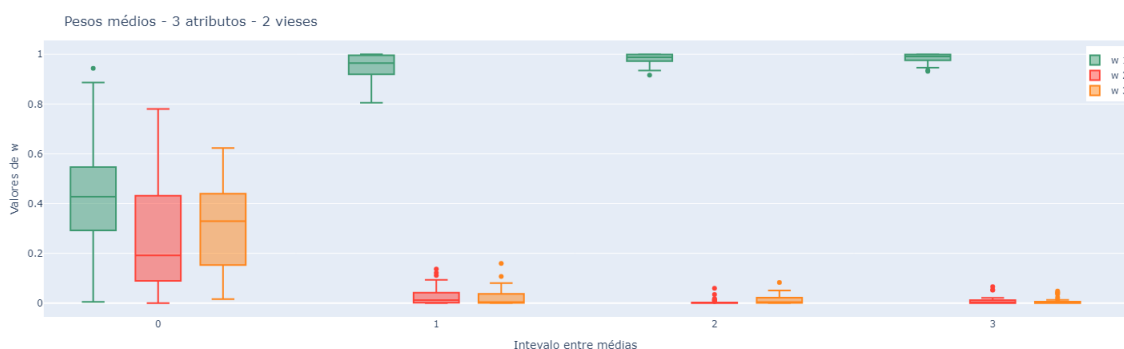


Figura 6: Box-plots de pesos para cada intervalo entre médias no caso de 3 critérios e 2 critérios enviesados.

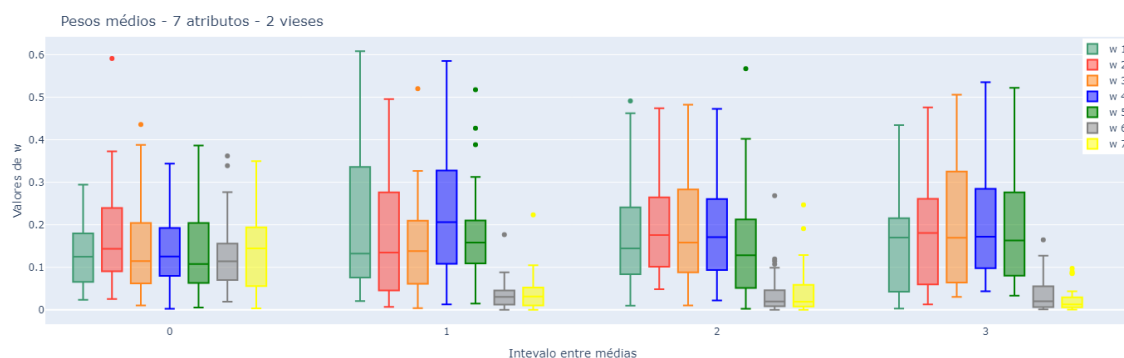


Figura 7: Box-plots de pesos para cada intervalo entre médias no caso de 7 critérios e 2 critérios enviesados.

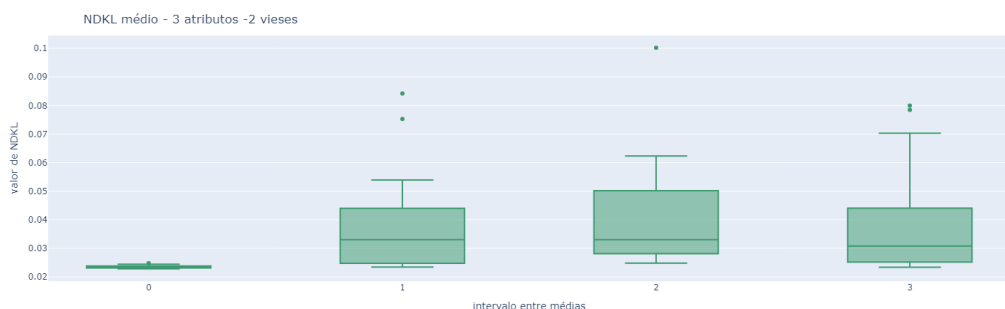


Figura 8: Box-plots de NDKL para cada intervalo entre médias no caso de 3 critérios e 2 critérios enviesados.

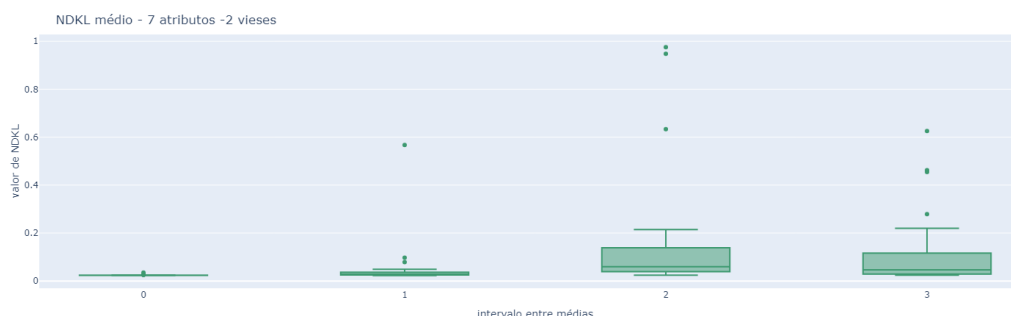


Figura 9: Box-plots de NDKL para cada intervalo entre médias no caso de 7 critérios e 2 critérios enviesados.

De modo geral, em ambos os casos, com 1 ou 2 critérios enviesados, em cenários que representam grupos pouco ou substancialmente desvalorizados, afetados por vieses, os resultados apontam não apenas uma tendência a 0, para desvalorização ou mesmo invalidação, dos pesos ligados aos critérios enviesados, como também evidencia a aleatoriedade dos demais pesos.

É importante acrescentar que foram feitos testes com conjuntos de dados com 5 e 10 critérios, cujos resultados seguem a mesma tendência que aqueles apresentados anteriormente para 3 e 7 critérios. No entanto, à medida que aumentava o número de critérios, para que houvesse conversão, foi necessário aumentar o número de iterações da Evolução Diferencial, aumentando o tempo de processamento demandando pelo algoritmo.

4.2. Segunda etapa: avaliação da otimização de NDKL com dados reais

Nesta etapa direcionamos os testes à otimização de NDKL em um conjunto de dados referentes a países, com o intuito de ordená-los de maneira equitativa. Tratam-se de dados de acesso livre, obtidos no *World Bank* [The World Bank Group, 2020]. Nele foram selecionados, para o ano de 2020, quatro critérios divididos entre indicadores econômicos e de sustentabilidade. São eles: área de floresta, expressa em porcentagem de terra; emissão de gás carbônico (CO_2), em toneladas de metros cúbicos por pessoa; expectativa de vida, em anos; e GNI per capita, Atlas method (*Gross National Income*), em dólares americanos.

Após exclusão de alternativas com dados faltantes, obtivemos um total de 179 alternativas que separamos em dois grupos: países emergentes e em desenvolvimento para o grupo sensível, com uma distribuição de 80.45%, e países avançados para o grupo não sensível, com uma distribuição de 19.55%. Para esta divisão de grupos, tomamos como base a lista de países avançados do Fundo Monetário Internacional (FMI) [International Monetary Fund, 2023], cuja estruturação leva em conta dados como PIB (Produto Interno Bruto), inflação, taxas de desemprego e crescimento empregatício, entre outros. Além disso, é importante acrescentar que utilizamos a técnica Min-Max para normalização dos dados e o valor resultante para a emissão de gás carbônico foi subtraído de 1, por ser um indicador de um efeito indesejável no âmbito da sustentabilidade.

A partir desses dados, os testes foram feitos em duas etapas. A primeira etapa envolve a busca dos pesos médios de cada critério, obtidos com a sequência de 30 repetições de otimizações de NDKL, cujos resultados ficaram entre 0.0150 e 0.0250, com uma média de 0.0249. Por conseguinte, podemos inferir, com o auxílio da Figura 10, que os pesos w_3 e w_4 , relacionados aos critérios 3 e 4, expectativa de vida e GNI, devem receber valores significativamente menores que os demais, w_1 e w_2 , quando o objetivo for a obtenção de ranqueamentos equitativos. Além disso, fica também evidente que o peso w_2 , ligado ao critério 2, emissão de CO_2 , é o que mais favorece países do grupo sensível. Isso pode estar atrelado ao fato de que países desenvolvidos, diferentemente dos demais países, geralmente, possuem uma grande produção industrial e, portanto, emitem mais CO_2 .

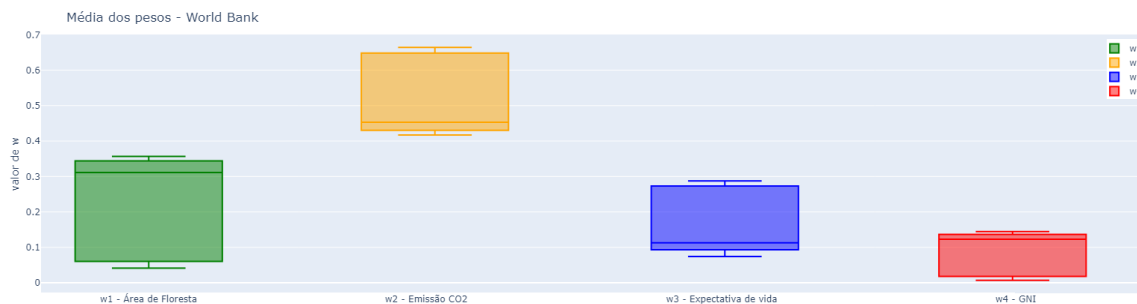


Figura 10: Box-plots de pesos para cada critério do conjunto de dados World Bank.

A segunda etapa do teste com dados reais consiste na comparação de diferentes ordenamentos dos países partindo de conjuntos de pesos semelhantes ou opostos à média obtida no processo de otimização, como mostra a Tabela 1. Na primeira linha desta tabela está representado um conjunto de pesos constantes, que não faz distinção entre os critérios, fazendo com que todos eles tenham igual impacto no cálculo da utilidade final de uma alternativa, a não ser pela influência dos próprios vieses, possivelmente existentes. Já na segunda linha está representado o conjunto de pesos equivalente aos pesos médios obtidos no processo de otimização, segundo o método que propomos, executado na etapa imediatamente anterior a esta. A terceira linha, por sua vez, apresenta um conjunto de pesos pautados na tendência observada no box-plot, sem necessariamente usar as médias, mas valorizando mais os dois primeiros critérios do conjunto de dados. E por último, a quarta linha apresenta um conjunto de pesos que faz o oposto daquele apresentado na linha anterior, com pesos que valorizam mais os dois últimos critérios do conjunto de dados, indo contra a tendência observada no box-plot.

Tabela 1: Conjuntos de pesos utilizados na segunda etapa de testes.

Conjuntos de pesos (w)	w ₁	w ₂	w ₃	w ₄
Pesos constantes	0,25	0,25	0,25	0,25
Pesos médios	0,32	0,45	0,11	0,12
Pesos similares	0,3	0,35	0,2	0,15
Pesos diferentes	0,2	0,2	0,3	0,3

A partir desses conjuntos de pesos, o conjunto de dados foi ordenado de diferentes formas e, para cada ordenamento, a equidade resultante foi calculada. Na Tabela 2, além do valor resultante de NDKL, podemos observar as porcentagens de alternativas do grupo sensível, países emergentes e em desenvolvimento (os quais representam 80.45% da população), presentes nos primeiros 10 (TOP 10), 15 (TOP 15) e 20 (TOP 20) itens dos ordenamentos.

Tabela 2: Resultados da segunda etapa de testes.

Resultados	Pesos constantes	Pesos médios	Pesos similares	Pesos diferentes
TOP 10	0%	80%	50%	0%
TOP 15	7%	80%	60%	0%
TOP 20	20%	85%	65%	0%
NDKL	0,819	0,025	0,313	0,997

A partir dessas informações, podemos observar que o conjunto de pesos, denominado

“Pesos diferentes”, cujos valores vão no sentido oposto à tendência encontrada no processo de otimização da medida de equidade, é aquele cujo ordenamento possui menos itens do grupo sensível nas primeiras posições, ainda que o grupo não sensível represente apenas 20% das alternativas. O NDKL resultante neste ordenamento é o maior e, portanto, menos equitativo de todos; nele o Brasil, por exemplo, ocupa a 40^a posição com uma utilidade total $u_{Brasil} = 0,53$. Também chama a atenção o fato de que um conjunto de pesos apenas similares aos pesos médios obtidos no processo de otimização reduziu em mais de 50% o valor de NDKL obtido com pesos constantes; nesse ordenamento o Brasil ocupa a 19^a posição com uma utilidade total $u_{Brasil} = 0,65$.

5. Conclusão

Neste trabalho, propusemos uma nova abordagem para lidar com problemas de ordenamento, no âmbito de MCDA, considerando premissas de equidade entre dois subgrupos de alternativas. Para tanto, sugerimos a otimização por Evolução Diferencial da medida NDKL em função dos pesos como um importante apoio ao processo de definição dos pesos dos critérios considerados nos cálculos das utilidades individuais das alternativas.

Para os experimentos com dados sintéticos, a partir das médias do peso w_j , onde j corresponde ao critério enviesado, observamos que a tendência à anulação deste peso é proporcional ao intervalo entre as médias dos grupos sensível (representantes de grupos marginalizados por vieses históricos) e não sensível (representantes de grupos privilegiados historicamente), ao passo que os demais pesos, quando não ligados ao critério enviesado, recebem valores aleatórios no intervalo $[0, 1]$. Isso indica que, ainda que não possa ser completamente eliminado dos dados, ou mesmo substituído, um critério com viés potencial, não deve ter sua importância sobreposta aos demais critérios. Esta afirmação é confirmada, neste estudo, pelos testes com dados reais que evidenciam a importância, em termos de equidade, da definição de pesos norteadas pelo processo de otimização, onde foi possível alcançar representatividade satisfatória do grupo sensível em uma seleção de países. Contudo, no caso de inviabilidade de implementação de um processo de otimização, a incorporação de alguma medida de equidade já trará mais consistência ao processo de ordenamento.

Para continuação deste estudo, evocando a necessidade de aumento do número de iterações para conversão do algoritmo no caso de dados sintéticos, os resultados médios de NDKL sugerem que a convergência do algoritmo apresenta um quadro mais instável à medida que cresce o número de critérios. Uma primeira perspectiva que surge a partir deste ponto, trata-se do estudo aprofundado da Evolução Diferencial, explorando os impactos e limites de cada um de seus parâmetros no processo de otimização em questão. Além disso, este estudo abre caminho e necessidade de exploração de casos com mais de um grupo sensível e a aplicação de outras técnicas de agregação.

Agradecimentos

O trabalho foi apoiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001; pelo Centro de Pesquisa Aplicada em Inteligência Artificial BIOS, processo nº 2020/09838-0; e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Agradecemos também ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e ao Programa de Pós-Graduação em Tecnologia (PPGT).

Referências

- Alvarez, P. A., Ishizaka, A., e Martínez, L. (2021). Multiple-criteria decision-making sorting methods: A survey. *Expert Systems With Applications*.
- Carvalho, D. V., Pereira, E. M., e Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*.

- Cem Geyik, S., Ambler, S., e Kenthapadi, K. (2019). Fairness-aware ranking in search & recommendation systems with application to linkedin talent search. *ACM SIGKDD Conference on Knowledge Discovery and DataMining (KDD'19)*.
- Dharmarathne, G., Bogahawaththa, M., McAfee, M., Rathnayake, U., e Meddage, D. (2024). On the diagnosis of chronic kidney disease using a machine learning-based interface with explainable artificial intelligence. *Intelligent Systems with Applications*, 22.
- International Monetary Fund, T. (2023). Web page. <https://www.imf.org/en/Publications/WEO/weo-database/2023/April/groups-and-aggregatesoae>. Acessado: 2020-05-22.
- Kullback, S. e Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., e Galstyan, A. (2021). A survey on bias and fairness in machine learnig. *ACM Computing Surveys*, 54.
- Nazemi, A. e Fabozzi, F. J. (2024). Interpretable machine learning for creditor recovery rates. *Journal of Banking and Finance*, 164.
- Price, K. V. e Storn, R. M. (1995). Differential evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization*.
- Price, K. V., Storn, R. M., e Lampinen, J. A. (2005). *Differential Evolution - Practical Approach to Global Optimization*. Springer, Berlim.
- Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*, volume 12. Springer.
- Roy, B. (2005). *Multiple Criteria Decision Analysis: State of the Art Surveys*, volume 78. Springer.
- Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*.
- Sandel, M. J. (2011). *Justiça: O que é fazer a coisa certa*, volume 35. Civilização Brasileira.
- Sang Ah Lee, M. e Singh, J. (2021). The landscape and gaps in open source fairness toolkits. *CHI Conference on Human Factors in Computing Systems (CHI '21)*.
- Tahery, S., Zahra Aftabi, S., e Farzi, S. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456.
- The World Bank Group, T. (2020). Web page. <https://data.worldbank.org/>. Acessado: 2020-05-15.
- Verma, S. e Rubin, J. (2018). Fairness definitions explained. *Fair-Ware'18: ACM/IEEE International Workshop on Software Fairness*.
- Yang, K. e Stoyanovich, J. (2017). Measuring fairness in ranked outputs. *Proceedings of SSDBM'17*.
- Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., e Baeza-Yates, R. (2017). Fa*ir : A fair top-k ranking algorithm. *CIKM'17*.