

Aprimorando Seleções Acadêmicas com Técnicas de Entropia e Método AHP-TOPSIS-2N

Enhancing Academic Selections with Entropy Techniques and AHP-TOPSIS-2N Method

Daniel Soares de Souza
<https://orcid.org/0000-0003-2210-5412>

Mestre em Gestão Pública. Instituto Federal de Educação, Ciência e Tecnologia de Brasília (IFB) – Brasil.
daniel.souza@ifb.edu.br.

Pedro Carvalho Brom
<https://orcid.org/0000-0002-1288-7695>

Mestre em Estatística. Instituto Federal de Educação, Ciência e Tecnologia de Brasília (IFB) – Brasil. pedro.brom@ifb.edu.br.

André Araújo Luchine
<https://orcid.org/0000-0002-8177-6390>

Bacharel em Ciências Econômicas. Universidade de Brasília (UnB) – Brasil. andreluchine@gmail.com.

Lucas Santos de Oliveira
<https://orcid.org/0000-0002-7128-8545>

Especialista em Estatística Aplicada. Instituto Federal de Educação, Ciência e Tecnologia de Brasília (IFB) – Brasil.
lucas.oliveira4@estudante.ifb.edu.br.

RESUMO

Este artigo explora a otimização de processos seletivos acadêmicos com o apoio de métodos tradicionais e técnicas de modelagem matemática e análise de dados. Através desses métodos e técnicas, esta pesquisa tem como objetivo desenvolver um processo que melhore a seleção de candidatos, respeitando suas singularidades e subjetividades. O estudo adota uma abordagem exploratória e aplicada, com foco na coleta de dados de processos seletivos de pós-graduação e na implementação de um modelo a ser replicado em outras seleções e em contextos diversos. A metodologia inclui a identificação de critérios relevantes usando o método de entropia, a raspagem automatizada de dados em documentos e a aplicação do método multicritério AHP-TOPSIS-2N. Ao combinar o processo de hierarquização do AHP e a técnica de classificação do TOPSIS, esse método multicritério permite uma análise detalhada e uma classificação justa das alternativas. Os resultados obtidos evidenciam a eficácia desses métodos em todo o processo de classificação dos candidatos, com uma avaliação holística e equilibrada. Com a análise de entropia, a qualidade da escrita nas cartas de intenção e nos currículos Lattes foram indicadores cruciais para a aprovação de candidatos, sublinhando a necessidade de critérios rigorosos de avaliação nestes aspectos. O estudo conclui que a integração de técnicas avançadas com métodos tradicionais permite uma avaliação mais holística e precisa dos candidatos, potencializando o processo seletivo em instituições de ensino superior.

Palavras-chave: Análise de Entropia. Método AHP-TOPSIS-2N. Modelagem Matemática. Processos Seletivos Acadêmicos.

ABSTRACT

This article explores the optimization of academic selection processes with the support of traditional methods and mathematical modeling and data analysis techniques. Through these methods and techniques, this research aims to develop a process that improves the selection of candidates, while respecting their singularities and subjectivities. The study adopts an exploratory and applied approach, focusing on collecting data from graduate selection processes and implementing a model to be replicated in other selections and

in different contexts. The methodology includes the identification of relevant criteria using the entropy method, automated data scraping from documents, and the application of multicriteria method AHP-TOPSIS-2N. By combining the AHP ranking process and the TOPSIS ranking technique, this multicriteria method allows for a detailed analysis and fair ranking of alternatives. The results obtained demonstrate the effectiveness of these methods throughout the candidate classification process, providing a holistic and balanced assessment. With the entropy analysis, the quality of writing in letters of intent and in Latte

CVs were crucial indicators for candidate approval, highlighting the need for rigorous evaluation criteria in these aspects. The study concludes that integration advanced techniques with traditional methods allows for a more holistic and accurate assessment of candidates, enhancing the selection process in higher education institutions.

Keywords: Academic Selection Processes. AHP-TOPSIS-2N Method. Entropy Analysis. Mathematical Modeling.

Recebido em 01/04/2024. Aprovado em 14/06/2024. Avaliado pelo sistema *double blind peer review*. Publicado conforme normas da APA.

<https://doi.org/10.22279/navus.v14.1894>

1 INTRODUÇÃO

No Brasil, as Instituições de Ensino Superior [IES] ofertam cursos que abrangem desde a graduação até a pós-graduação. Em particular, a pós-graduação é oferecida em duas categorias: a *Lato Sensu* e a *Stricto Sensu*. Essas categorias compreendem os programas de aperfeiçoamento, doutorado, especialização, mestrado e outros (Lei n. 9.394, 1996). Ao contrário dos cursos de graduação, cada instituição possui seus próprios critérios e processos seletivos para o ingresso de estudantes em cursos de pós-graduação (Soares et al., 2011).

Com o objetivo de garantir a seleção dos candidatos mais adequados para os cursos, os processos seletivos incluem etapas como, por exemplo, análise de documentação de inscrição, análise de histórico escolar, avaliação curricular, análise do projeto, arguição oral do projeto, carta de recomendação, cartas de apresentação/intenção ou plano de trabalho, certificado de proficiência em língua estrangeira, entrevista, prova de conhecimentos específicos, prova de redação e resultado de teste de associações representativas dos cursos e carreiras (Soares et al., 2011; Silva & Silva, 2016; Cunha, 1974). Nesse cenário, é razoável inferir que os processos seletivos concentram uma quantidade significativa de informações acadêmicas e profissionais de seus respectivos candidatos.

A partir dessas informações, são identificados e escolhidos os candidatos mais adequados e qualificados para as vagas disponíveis. Contudo, através de análise preditiva, existe a possibilidade de construir um modelo que faz previsões com base em padrões extraídos de dados históricos (Fernandes & Rêgo, 2023). No contexto acadêmico, os modelos preditivos são geralmente utilizados para prever a evasão escolar e o desempenho dos estudantes em disciplinas específicas (Funchal et al., 2016; Souza, 2021; Morais et al., 2023; Oliveira & Medeiros, 2024).

Em relação aos processos seletivos, os algoritmos do modelo preditivo podem, através de Aprendizado de Máquina, ser utilizados para analisar automaticamente ensaios de admissão, recomendações e outros dados de interesse (Shahiri et al., 2015). Nesse cenário, os processos como coleta de dados, análise preliminar e comunicação com candidatos também podem ser automatizados com conjunto de instruções de códigos de programação (Romero et al., 2008).

Ao utilizar técnicas de processamento de texto nas etapas de coleta, processamento e avaliação de dados dos candidatos, o desenvolvimento desse processo automatizado/semiautomatizado possui um impacto significativo dentro do contexto seletivo-acadêmico, visto que essas técnicas permitem analisar de forma eficiente as informações contidas nas cartas de intenção e currículos *Lattes* dos candidatos. Além disso, a integração do modelo através de entropia permite extrair informações relevantes e convertê-las em uma matriz de decisão. Contudo, é importante mencionar que essa abordagem levanta questões éticas e práticas, como a privacidade dos dados, potenciais vieses nos algoritmos, entre outros (Fernandes & Rêgo, 2023).

Em relação ao processo seletivo, também existem diversas ferramentas multicritério que podem auxiliar na avaliação e classificação dos candidatos, com base em múltiplos critérios (Behzadian et al., 2012). Nesse caso, a análise multicritério é uma estratégia para avaliar os candidatos de uma maneira metodológica que considere diferentes critérios relevantes para o sucesso acadêmico e profissional. Contudo, a implementação dessas ferramentas requer uma compreensão sólida de sua teoria e prática.

Essa abordagem afeta diretamente a eficiência do processo seletivo e, conseqüentemente, promove uma avaliação mais justa e transparente. Dessa forma, a delimitação desse estudo visa aprofundar a compreensão das técnicas e ferramentas utilizadas no processo de seleção acadêmica, contribuindo para aprimorar a qualidade e a eficácia desse processo em IES.

Embora o recrutamento de candidatos a cursos acadêmicos possua técnicas e processos seletivos específicos, verifica-se a necessidade de adequação dos critérios de seleção às demandas e características particulares de cada instituição, assim como à disponibilidade e acessibilidade dos dados dos candidatos. Além disso, devido à complexidade e diversidade dos currículos acadêmicos, é cada vez mais necessário o uso de um sistema automatizado capaz de analisar e classificar, de forma precisa e eficiente, as informações relevantes. Portanto, o estudo se limita a explorar, em um contexto específico, como a aplicação das técnicas apresentadas pode contribuir para aprimorar o processo de seleção acadêmica, considerando as particularidades e desafios enfrentados pelas instituições.

Para mapear o atual estado de conhecimento na literatura nacional, foi realizado um levantamento no Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior [CAPES] a partir dos seguintes descritores: "Processo Seletivo" e "Pós-Graduação". Logo após, os resultados foram filtrados para direcionar somente artigos publicados entre 2020 e 2025. Com esse procedimento, foram encontrados 25 artigos.

Após a leitura dos resumos de cada um dos artigos encontrados, observa-se que os estudos abordam os seguintes temas: 1) panorama das políticas de inclusão e de ações afirmativas; 2) elaboração dos instrumentos nos processos seletivos; 3) dificuldades de inserção dos candidatos nos programas; 4) identificação de critérios influenciadores na escolha dos candidatos para um programa; 5) criação de cursos preparatórios para os candidatos ao processo seletivo acadêmico; 6) análise do estado atual de elaboração/funcionamento dos programas de pós-graduação.

A lacuna a ser preenchida, portanto, reside na ausência de estudos relacionados ao uso de técnicas avançadas de modelagem matemática e automatização no contexto de processos seletivos acadêmicos (em especial IES no Brasil). Isso sugere que a pesquisa realizada aqui se encontra na fronteira do conhecimento.

Mediante este contexto, emerge a seguinte questão de pesquisa: Como é possível incrementar estratégias que visam a melhoria de processos seletivos de cursos acadêmicos, levando em consideração o recrutamento, o processamento e a avaliação sistematizada de dados dos candidatos, de maneira que sejam tratadas as suas subjetividades naturais?

O objetivo deste artigo é desenvolver, a partir da integração de técnicas de modelagem matemática em processos seletivos acadêmicos, incrementos de automatização ou semi-automatização para que se aprimore, dado o recrutamento, a avaliação de dados dos candidatos, enquanto contempla e respeita suas singularidades e subjetividades. De maneira específica ao cumprimento geral do objetivo, têm-se: a) elencar de maneira fundamentada os principais critérios que devem estar presentes nos processos seletivos; b) implementar modelos matemáticos adaptáveis que facilitem a coleta, processamento e interpretação de dados dos candidatos, no intuito de uma avaliação mais holística e precisa; c) Avaliar a eficácia do modelo em termos de eficiência, precisão e equidade nas decisões tomadas para classificar e selecionar os candidatos, com base em critérios normalizados e ponderados, consoante o grau de importância no processo seletivo.

Diante disto, as especificidades da investigação consistiram em identificar os principais critérios e subjetividades presentes nos processos seletivos de cursos acadêmicos e, como estudo de caso, aplicou-se técnicas e métodos em um curso de pós-graduação no Instituto Federal de Brasília. Além disso, foram exploradas técnicas para tratamento de medidas de incerteza ou desordem aparente, sendo estas adequadas para o tratamento e análise de dados dos candidatos. Ao aplicar tal procedimento, foi possível considerar tanto critérios objetivos quanto subjetivos e, por fim, se avaliou a eficácia e precisão de uma sistemática padrão em comparação com os processos seletivos tradicionais. Com esses direcionamentos em mente, buscou-se contribuir para o aprimoramento dos processos seletivos, gerados em uma seleção mais eficiente, capaz de enfrentar a diversidade, como também controlar excessos de eventuais vieses de seleção.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Processos seletivos acadêmicos

No Brasil, a qualidade das IES tem sido uma preocupação constante, refletida na atuação de órgãos governamentais como a CAPES e o Instituto Nacional de Estudos e Pesquisas Educacionais [INEP]. Esses órgãos desempenham um papel fundamental na avaliação e acompanhamento do desempenho das instituições de graduação e pós-graduação (Igarashi et al., 2008). No âmbito da pós-graduação, em particular, o sistema de avaliação da CAPES é amplamente reconhecido como eficiente, devido aos critérios que incluem a proposta do programa, a avaliação do corpo docente e discente, a produção intelectual e a inserção social (CAPES, 2010; Leite et al., 2000).

A seleção dos candidatos para os cursos acadêmicos desempenha um papel que, de certa maneira, influencia na avaliação desses cursos, visto que a avaliação da CAPES atribui 35% de peso à seleção dos candidatos (CAPES, 2010). Contudo, dentro do escopo da seleção acadêmica, existe uma dicotomia entre critérios objetivos e subjetivos que guiam o processo de avaliação dos candidatos.

Santos e Machado (2018) definem os critérios objetivos como métricas quantitativas. Estes incluem o histórico escolar ou notas, visto que esses instrumentos fornecem uma medida quantitativa direta do desempenho acadêmico anterior do candidato. Além disso, as pontuações em testes padronizados, como *Scholastic Assessment Test* [SAT], *Graduate Record Examination* [GRE] e *Graduate Management Admission Test* [GMAT], atuam como indicadores do potencial acadêmico em contextos específicos. Não menos importantes são os projetos de pesquisa anteriores, como publicações ou participações em conferências, e a experiência prévia relacionada à área de estudo, como estágios ou trabalhos voluntários.

Por outro lado, os critérios subjetivos oferecem uma visão mais profunda da personalidade, motivações e caráter de um candidato. Souza et al. (2014) evidenciam a importância das cartas de recomendação no processo de admissão, pois oferecem perspectivas valiosas de professores, supervisores ou colegas sobre a capacidade e caráter do candidato. Igualmente importantes são os ensaios ou declarações pessoais, que permitem aos candidatos expressar suas motivações, interesses e visões. Além disso, entrevistas, muitas vezes conduzidas por painéis acadêmicos, servem como uma ferramenta para avaliar diretamente habilidades de comunicação, paixão e adequação ao programa ou

instituição. O envolvimento em atividades extracurriculares também pode fornecer percepções sobre a personalidade e interesses de um candidato.

No entanto, ao avaliar os candidatos com base nesses critérios, surgem preocupações sobre possíveis vieses inerentes. Nickerson (1998) identifica o viés de confirmação como um dos mais prevalentes, onde os avaliadores tendem a dar mais peso às informações que confirmam suas crenças ou ideias preexistentes. Além disso, o Efeito Halo pode causar distorções na percepção positiva em um critério e, por consequência, pode influenciar indevidamente a avaliação de outros critérios (Kahneman, 2012). O Efeito Halo é um termo descrito pelo psicólogo americano Edward Thorndike para descrever o potencial que o cérebro humano tem de analisar, julgar, concluir e definir uma pessoa a partir de uma única característica (Thorndike, 1920).

Com o objetivo de tratar e não eliminar a subjetividade existente em problemas desta natureza, pode-se investigar o emprego desta abordagem através de um experimento ilustrativo. Neste contexto, observa-se que, em muitos casos, a tomada de decisão em amplo espectro ocorre efetivamente em função de opiniões coletadas junto a pessoas conhecidas e/ou que trabalham na área ou atuam diretamente na situação problema. Isso infere um alto grau de subjetividade no tratamento das informações quanto à classificação e ordenamento e, conseqüentemente, resultarão na escolha mais assertiva do que quer que seja o objeto da decisão a ser tomada.

Sob a perspectiva de Trevizano e Freitas (2005), a abordagem proposta não elimina a subjetividade inerente aos problemas decisórios, apenas busca tratar destes através do emprego de um método que possibilita a verificação da consistência dos julgamentos. Em particular, tais julgamentos preferencialmente devem ser conduzidos por profissionais que possuem conhecimentos na área do problema em questão, como o próprio corpo docente e áreas afins. Este é o principal motivo pelo qual não se recomenda o emprego de uma abordagem tão técnica e de considerável complexidade para pessoas consideradas leigas.

Os julgamentos de preferência das alternativas à luz de cada critério e da importância dos critérios em relação ao foco principal são obtidos a partir da opinião de especialistas, não se caracterizando como uma verdade, visto que existe um alto grau de subjetividade nestes julgamentos. Mais precisamente, os resultados obtidos caracterizam-se pela opinião única e pessoal do respectivo avaliador, não devendo ser generalizadamente expandida para decisões de todo e qualquer processo seletivo (Trevizano & Freitas, 2005).

2.2 Modelagem matemática e automatização como ferramentas de equidade

A modelagem matemática é o processo de usar linguagem matemática para descrever e analisar problemas do mundo real (Triantaphyllou, 2000). Ao fazer isso, os pesquisadores podem transformar problemas complexos e muitas vezes ambíguos em estruturas compreensíveis, permitindo soluções objetivas. No contexto da seleção acadêmica, isso se manifesta na forma de decisões multicritério, onde vários critérios (tanto objetivos quanto subjetivos) são ponderados e avaliados simultaneamente para chegar a uma decisão final sobre a admissão de um candidato.

A tomada de decisões em processos seletivos acadêmicos requer a identificação precisa de critérios que refletem os valores, objetivos e missões de uma instituição educacional. Estes critérios, muitas vezes, são múltiplos e variados, tornando a seleção uma tarefa complexa. A utilização

do método de entropia surge como uma abordagem promissora para determinar a relevância de cada critério, otimizando a decisão final (Shannon, 1948; Zeleny, 1982).

Originado da teoria da informação, o conceito de entropia é uma medida da magnitude de ambiguidade ou caos em um conjunto de dados (Shannon, 1948). Em outras palavras, a entropia é uma medida de incerteza ou desordem em um conjunto de dados. Em termos de classificação, quando a entropia é alta, significa que temos muita incerteza sobre a classificação; quando é baixa, temos mais certeza (Zeleny, 1982).

No contexto dos processos seletivos, a entropia auxilia na determinação de quais critérios possuem maior variação ou dispersão de informações entre os candidatos, sinalizando seu potencial relevante no processo. Critérios com alta entropia indicam uma grande diversidade nas respostas ou desempenho dos candidatos, tornando-os essenciais para a diferenciação e decisão. Por outro lado, critérios com baixa entropia têm respostas semelhantes entre os candidatos e, talvez, possam ser menos determinantes no processo de seleção.

A utilização do método de entropia proporciona uma filtragem ao reduzir a dimensão da informação, primando pela importância relativa dos critérios no processo seletivo. Ao adotar essa abordagem, as instituições podem criar um sistema de avaliação mais impessoal, imparcial e adaptado à sua missão educacional (Aguiar et al., 2015).

Em termos matemáticos, a entropia de uma variável aleatória Y é calculada usando a fórmula:

$$H(Y) = - \sum_{y \in Y} p(y) \log_2 p(y)$$

Na equação acima, $p(y)$ denota a probabilidade de um resultado específico y (aprovado ou não aprovado).

Por outro lado, a entropia condicional é uma medida importante na análise de variáveis aleatórias (Shannon, 1948). Em termos matemáticos, a entropia condicional $H(Y|X)$ de Y dado X é definida como:

$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X = x)$$

onde $H(Y|X = x)$ representa a entropia de Y com a condição de X fixado em um valor específico x , e $p(x)$ é a probabilidade de X assumir esse valor.

Em relação ao processo seletivo acadêmico, a entropia pode ser combinada com outras técnicas multicritério para uma análise ainda mais refinada. Por exemplo, a entropia pode ser usada em conjunto com o Método *Analytic Hierarchy Process* [AHP] ou o Método *Preference Ranking Organization METHod for Enrichment Evaluation* [PROMETHEE] para aprimorar a ponderação de critérios e a tomada de decisão (Neeraj et al., 2019). Contudo, o método AHP-TOPSIS-2N representa uma abordagem avançada na tomada de decisão, particularmente adequada para situações em que os aspectos subjetivos desempenham um papel significativo.

O método AHP-TOPSIS-2N é indicado para cenários onde as preferências pessoais e as percepções individuais influenciam fortemente a escolha entre diferentes alternativas e critérios. Como destacado por Saaty (2008), o método AHP é eficaz na quantificação de elementos subjetivos, transformando raciocínios qualitativos em quantitativos. Essa capacidade é apropriada para

contextos onde as métricas são intrinsecamente subjetivas e complexas, como a avaliação da satisfação do cliente ou a definição de prioridades estratégicas.

Além disso, o método AHP-TOPSIS-2N é particularmente útil em situações com um grande número de alternativas, dado que não possuem restrições quanto ao número de opções a serem avaliadas. Esta característica o torna adequado para ambientes de decisão complexos, onde diversas opções precisam ser consideradas. Segundo Hwang e Yoon (1981), o método TOPSIS, do inglês *Technique for Order of Preference by Similarity to Ideal Solution*, é eficiente na análise de múltiplas alternativas, facilitando a ordenação em um conjunto extenso.

No entanto, ao utilizar critérios quantitativos, é importante observar as limitações de até 15 critérios para garantir a previsão dos cálculos de consistência, conforme especificado por Saaty (1980). Esta limitação não reflete uma restrição metodológica, mas sim uma adaptação prática para manter a precisão e a confiabilidade do processo de tomada de decisão. A gestão da tolerância à inconsistência, conforme apontado por Saaty (1980), não é uma falha do método, mas uma representação da complexidade e das nuances da realidade enfrentada durante o processo decisório.

Na aplicação do método AHP-TOPSIS-2N, o primeiro passo envolve a utilização do AHP para determinar os pesos dos critérios. Este processo se inicia com a criação de uma matriz A de comparação par a par para os critérios, onde os elementos a_{ij} representam a importância relativa do critério i em relação ao critério j . Em termos matemáticos, a matriz A é definida da seguinte forma:

$$A = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \dots & 1 \end{pmatrix}$$

Os pesos dos critérios são derivados a partir desta matriz A , normalmente através do cálculo do vetor de prioridades ou autovalores.

Com os pesos dos critérios estabelecidos, o TOPSIS é aplicado para classificar as alternativas. Inicialmente, constrói-se uma matriz de decisão normalizada R usando a fórmula:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^n x_{ij}^2}}$$

Na igualdade acima, x_{ij} é o valor do critério i para a alternativa j , e n é o número total de alternativas. Em seguida, calcula-se a matriz de decisão ponderada V através da seguinte expressão matemática:

$$v_{ij} = w_i \cdot r_{ij}$$

Na equação acima, w_i é o peso do critério i . A solução ideal positiva A^+ e a solução ideal negativa A^- são determinadas:

$$A^+ = \{\max_j v_{ij}; i \in \text{critérios de benefício}, \min_j v_{ij}; i \in \text{critérios de custo}\}$$

$$A^- = \{\min_j v_{ij}; i \in \text{critérios de benefício}, \max_j v_{ij}; i \in \text{critérios de custo}\}$$

As distâncias de cada alternativa até as soluções ideais positiva e negativa são calculadas com as seguintes fórmulas:

$$D_j^+ = \sqrt{\sum_{i=1}^m (v_{ij} - A_i^+)^2}$$

$$D_j^- = \sqrt{\sum_{i=1}^m (v_{ij} - A_i^-)^2}$$

Finalmente, calcula-se o score de similaridade para cada alternativa:

$$S_j = \frac{D_j^-}{D_j^+ + D_j^-}$$

As alternativas são então classificadas com base em S_j , com scores mais altos que indicam uma posição mais elevada na classificação.

Para facilitar a aplicação do método AHP-TOPSIS-2N, Bozza et al. (2020) criou o programa *Three Decision Methods* [3DM], um ambiente *online* eficiente e acessível para a aplicação desta abordagem de tomada de decisão multicritério. Esse programa permite aos usuários inserir dados, realizar análises e visualizar resultados de forma intuitiva e simplificada, tornando o processo de tomada de decisão mais acessível, especialmente para aqueles sem conhecimento profundo em programação ou análise estatística.

A funcionalidade central do 3DM reside na sua capacidade de analisar grandes conjuntos de dados, aplicando os métodos AHP e TOPSIS sequencialmente para gerar classificações e avaliações planejadas. Além disso, o 3DM se destaca pela adaptabilidade, pois esse programa é aplicável a uma variedade de contextos de tomada de decisão, desde análises empresariais até avaliações acadêmicas. A capacidade de analisar dados de maneira eficiente e precisa torna o *software* uma ferramenta valiosa para pesquisadores, gestores e tomadores de decisão (Bozza et al., 2020).

2.3 Incrementos de instrumentos para melhoria de processos

Com a ascensão da tecnologia, a automatização e semi-automatização tornaram-se elementos cruciais nos processos seletivos (Brynjolfsson & Hitt, 2000). A automação pode lidar com tarefas repetitivas e de alta frequência, como a capacidade de manter um pelotão veicular estável em diferentes topologias de comunicação (Oliveira et al., 2020). Isso não apenas reduz o tempo, mas também garante uma consistência que pode ser difícil de alcançar em avaliações manuais.

No contexto de seleção acadêmica, a raspagem de dados pode lidar com os documentos submetidos pelos candidatos. Também conhecida como *web scraping*, a raspagem de dados envolve a extração automatizada de informações

de documentos em formato *Portable Document Format* [PDF] utilizando *scripts* em linguagem *Python* que são capazes de decodificar e analisar documentos digitais estruturados de maneira complexa, como tabelas, gráficos e textos formatados (Munzert et al., 2014).

Mediante o uso de linguagem *Python*, a automação dessa raspagem de dados não apenas aumenta a eficiência e a amplitude da coleta de dados, mas também permite uma maior personalização no processo de extração ao adaptar-se às especificidades dos documentos em PDF. Por isso, essa técnica é utilizada em pesquisas empíricas e meta-análises, visto que se torna possível realizar análises mais profundas e abrangentes (Figueiredo Filho et al., 2014; Cassiano et al., 2016).

Contudo, a raspagem de dados de PDFs apresenta alguns desafios, como a variabilidade na formatação dos documentos. Além das questões técnicas, há considerações éticas e legais relacionadas à extração de dados sem autorização, que devem ser cuidadosamente avaliadas para garantir a conformidade com as normas de direitos autorais e privacidade (Domingues & Domingues, 2023).

Além disso, um dos desafios dos processos de automatização é a possibilidade de codificar vieses existentes nas ferramentas automatizadas, levando a decisões que refletem as desigualdades do passado. Além disso, a automação pode não ser adequada de capturar nuances ou qualidades intangíveis, evidentes em uma revisão manual.

Na etapa de modelagem matemática são empregados um conjunto de procedimentos que objetivam a construção de um paralelo que explique, matematicamente, os fenômenos presentes no cotidiano dos indivíduos, auxiliando-os a fazer previsões e tomar decisões (Burak, 1992). Quando corretamente aplicada, a modelagem matemática tem o potencial de identificar e mitigar vieses inerentes. Para ilustrar, técnicas de otimização podem ser aplicadas para garantir que uma variedade de critérios seja considerada de forma equitativa, evitando a super-representação de um único critério dominante.

Quando utilizada de maneira judiciosa, a automatização pode também desempenhar um papel na promoção da equidade. Sistemas automatizados podem ser programados para não considerar fatores potencialmente enviesados, como origem étnica ou socioeconômica, garantindo que as decisões sejam baseadas apenas em méritos acadêmicos e outros critérios relevantes (Nobre, 2020; Rocha et al., 2020; Takakura & Duarte, 2022). Portanto, a integração da modelagem matemática e automação em processos seletivos acadêmicos oferece uma oportunidade para instituições melhorarem a eficiência e a equidade de suas decisões de admissão. Entretanto, recomenda-se abordar essa integração com cuidado e consciência para maximizar seus benefícios e minimizar potenciais desafios.

Ao propor um modelo de automação em processos seletivos acadêmicos, aspectos como eficiência, precisão e equidade precisam ser considerados para garantir sua adequação ao contexto acadêmico.

A eficiência em seleções acadêmicas refere-se à capacidade de um sistema ou processo de realizar suas tarefas de seleção com o mínimo de desperdício de recursos e tempo (Chiavenato, 2009). Isso envolve a análise rápida e correta de candidaturas, bem como a tomada de decisões de admissão em tempo hábil (Chiavenato, 2009). A precisão, por outro lado, foca na correção das decisões tomadas. Um processo preciso identifica e admite candidatos possivelmente qualificados e adequados para o programa ou curso em questão (Chiavenato, 2009). A equidade aborda a justiça e imparcialidade do processo

de seleção, garantindo que todos os candidatos tenham a mesma oportunidade (Barreto et al., 2023).

Para avaliar a eficiência de um modelo, pode-se utilizar métricas como tempo de processamento (Diverio et al., 1995). Para avaliar a qualidade preditiva do modelo, foram utilizadas métricas como Acurácia, *Recall* e Precisão (Strauss et al., 2023). Em relação à equidade, o modelo pode ser avaliado através da análise de grupos desfavorecidos (Rosa & Gonçalves, 2015).

A adoção de modelos matemáticos no processo de seleção acadêmica tem o potencial de aumentar a objetividade e reduzir vieses humanos. No entanto, esses modelos são tão bons quanto os dados e os critérios em que se baseiam. Se os dados históricos usados para treinar um modelo contêm vieses, o modelo também os terá (Barocas & Selbst, 2016).

Dito isso, técnicas modernas de modelagem oferecem maneiras de identificar e controlar vieses. Ao combinar dados robustos, análise cuidadosa e revisão regular, é possível desenvolver modelos que promovam a justiça e a equidade na seleção acadêmica, garantindo que todos os candidatos sejam avaliados com base em seus méritos e qualificações (Corbett-Davies et al., 2017).

3 METODOLOGIA

Este artigo consiste em pesquisa exploratória e aplicada, conforme a classificação proposta por Gil (2010). As pesquisas exploratórias têm como objetivo proporcionar maior familiaridade com o problema, tornando-o explícito ou construindo hipóteses. Ao considerar a natureza inovadora da proposta que consiste na integração da técnica de entropia, na automatização para raspagem de dados e no método multicritério em processos seletivos, é essencial uma abordagem exploratória que permite o aprofundamento e a compreensão das variáveis envolvidas. Já o caráter aplicado da pesquisa corre da sua intenção prática de aperfeiçoar e otimizar processos seletivos reais em contextos acadêmicos.

A princípio, foram coletados dados do processo seletivo de um curso de pós-graduação ofertado pelo Instituto Federal de Brasília, objeto do estudo de caso desta pesquisa. Esse processo foi composto de duas etapas: 1) Análise curricular; 2) Carta de intenção. Na análise curricular, os candidatos preenchem uma tabela com as suas respectivas experiências profissionais, participação em eventos, apresentação de trabalho e produção acadêmica. Na etapa seguinte, a carta de intenção apresenta a trajetória profissional/acadêmica e o interesse de pesquisa do candidato, com um mínimo de 800 e um máximo de 1000 palavras. Além disso, nesse processo, o candidato também disponibiliza dados sensíveis como, por exemplo, o endereço eletrônico do seu respectivo currículo *Lattes*. De maneira geral, esses dados contêm informações sobre os documentos comprobatórios exigidos no próprio processo e, por isso, foram resguardados os devidos cuidados aos dados sensíveis dos candidatos, conforme normas estabelecidas na Lei n. 13.709 (2018).

A seguir, estão detalhadas as etapas metodológicas adotadas para o desenvolvimento deste artigo. Essas etapas são compostas de três fases principais, cada uma dedicada a uma etapa específica do processo de seleção: coleta de dados, identificação de critérios e avaliação multicritério.

3.1 Modelagem e estruturação técnica

Na primeira etapa, foi realizada uma raspagem de dados nos arquivos em formato PDF do processo seletivo deste estudo por meio de linguagem de programação *Python*. Foi adotado o uso de linguagem de programação *Python* para raspagem de PDFs devido à sua flexibilidade e às poderosas bibliotecas disponíveis, como *PyPDF2* e *PDFMiner*. Essas ferramentas transformam os dados extraídos em formatos analisáveis e, por consequência, facilitam a coleta de informações de uma variedade de fontes documentais, especialmente útil em campos como a pesquisa acadêmica, onde os PDFs são um meio comum de comunicação de dados (Boettiger, 2015).

A depuração dos dados foi realizada por meio de raspagem de dados, onde cada ocorrência identificada em cada critério foi contabilizada. Essa abordagem permitiu uma análise quantitativa mais precisa e uma avaliação qualitativa mais fundamentada.

Após a raspagem dos dados, foram identificados os critérios mais relevantes para o processo seletivo. Para isso, foi utilizado o método de entropia, pois esse método permite determinar a importância relativa de cada critério com base na dispersão das informações. Nesta pesquisa, a entropia condicional quantificou a incerteza na classificação dos candidatos, aprovado ou não, dada a informação extraída dos documentos fornecidos (cartas de intenção e currículos *Lattes*).

Na carta de intenção, a entropia ajudou a discernir a clareza e a coerência das motivações e objetivos do candidato, assim como o alinhamento destes com os objetivos do programa. Por outro lado, no currículo *Lattes*, a entropia quantificou a relevância e a diversidade das experiências acadêmicas e profissionais, avaliando a profundidade e a amplitude das realizações dos candidatos.

A análise revelou que as variáveis "Qualidade da Escrita" contida na carta de intenção e "Qualidade da Escrita" contida no currículo *Lattes* possuem as menores entropias condicionais, indicando que esses elementos são fortes indicadores para determinar a aprovação de um candidato. A baixa entropia condicional sugere que estas variáveis são fortes indicadores do resultado final de aprovação ou não aprovação, proporcionando uma base sólida para a tomada de decisão. Isso implica que a qualidade da escrita, tanto na carta de intenção quanto no currículo *Lattes*, deve ser um critério chave no processo de seleção. Esta conclusão é suportada pela entropia condicional combinada de zero para pares como "Qualidade da Escrita" da carta de intenção e "Qualificações" do currículo *Lattes* e "Qualidade da Escrita" na carta de intenção e "Qualidade da Escrita" no currículo *Lattes*. Tais combinações indicam uma forte interdependência entre estas variáveis, reforçando ainda mais a importância da qualidade da escrita na seleção dos candidatos.

Portanto, para a tomada de decisão no processo seletivo, recomenda-se a implementação de critérios rigorosos de avaliação da qualidade da escrita nas cartas de intenção e nos currículos *Lattes* dos candidatos. No contexto do presente estudo, a aplicação da análise de entropia revelou uma sinalização sobre as variáveis mais informativas no processo de seleção.

A Tabela 1 resume as entropias condicionais das variáveis individuais e das combinações de variáveis, oferecendo uma visão quantitativa da importância relativa de cada uma delas no contexto da decisão de aprovação de candidatos.

Tabela 1

Quadro-resumo da análise de entropia para avaliação de processo seletivo

| Categoria | Variável/Combinação de variáveis | Entropia condicional |
|------------------------------|--|-----------------------------|
| Variáveis individuais | Qualidade da Escrita (CI) | 0.097 |
| | Qualidade da Escrita (CL) | 0.099 |
| Combinação de duas variáveis | Qualidade da Escrita (CI) e Qualificações (CL) | 0.000 |
| | Qualidade da Escrita (CI) e Qualidade da Escrita (CL) | 0.000 |
| | Experiência Prévia (CI) e Qualidade da Escrita (CI) | 0.000 |
| | Objetivos Profissionais (CI) e Qualidade da Escrita (CI) | 0.032 |
| | Objetivos Profissionais (CI) e Qualidade da Escrita (CL) | 0.032 |

Nota: CI = Carta de Intenção; CL = Currículo *Lattes*.

Para aplicar o método AHP-TOPSIS-2N na avaliação do processo seletivo, foi desenvolvida uma matriz de decisão que ressalta a relevância das tradições de critérios propostos pela AHP. Nesta matriz, apresentada na Tabela 2, cada candidato é avaliado com base em critérios específicos, como 'Experiência Acadêmica e Publicações', 'Qualificações Profissionais e Desempenho Acadêmico', entre outras interações pertinentes. Os valores numéricos atribuídos a esses critérios, especificados pela ocorrência computada dos respectivos elementos, refletem uma avaliação precisa e específica de cada candidato.

Tabela 2

Estrutura da matriz de decisão formatada

| ID Candidato | Qualidade Escrita e Qualificações | Qualidade Escrita Carta e Lattes | Experiência e Qualidade Escrita | Objetivos e Qualidade Escrita Carta | Objetivos e Qualidade Escrita Lattes |
|---------------------|--|---|--|--|---|
| Candidato 1 | Valor numérico | Valor numérico | Valor numérico | Valor numérico | Valor numérico |
| Candidato 2 | Valor numérico | Valor numérico | Valor numérico | Valor numérico | Valor numérico |
| : | : | : | : | : | : |
| Candidato <i>N</i> | Valor numérico | Valor numérico | Valor numérico | Valor numérico | Valor numérico |

O processo de tomada de decisão para a seleção dos candidatos foi estruturado utilizando o método AHP-TOPSIS-2N, com uma abordagem específica que permite a diferenciação e ponderação classificada entre critérios qualitativos e quantitativos. A modelagem do problema envolveu várias etapas fundamentais para garantir a robustez e a eficácia da análise. Inicialmente, o AHP foi aplicado para determinar os pesos dos critérios, considerando os aspectos quantitativos e qualitativos. Em seguida, o método TOPSIS foi utilizado para classificar os candidatos com base nos critérios ponderados.

A entropia, neste caso, serviu como uma medida da distribuição e da variação das informações relativas aos candidatos em cada critério. Ao captar a frequência de ocorrência de cada candidato em cada critério, a entropia condicional permitiu identificar padrões de concentração ou dispersão. Isso pode ser particularmente útil para identificar áreas onde os candidatos são fortes ou fracos em comparação uns com os outros.

A etapa seguinte consiste na aplicação do método AHP-TOPSIS-2N, utilizando duas normalizações diferentes que produzem os resultados. Por se tratarem de critérios quantitativos, os dados apurados para cada candidato foram inseridos, adotando-se a técnica de avaliação de desempenho de alternativas através da similaridade da mesma com uma solução ideal e também levando em conta a avaliação dos critérios segundo a Escala de Saaty pelos responsáveis pelas decisões. De acordo com essa técnica, a melhor alternativa é aquela mais próxima da solução ideal e mais distante da solução não ideal (Bozza et al., 2020). Neste contexto, os critérios foram tratados como monotônicos de vantagem, onde valores maiores indicam uma preferência maior.

Na etapa final, foi realizada a indicação das prioridades entre os critérios. Para ajustar as prioridades no AHP, cada par de critérios deve ser reavaliado com base na sua importância relativa ao objetivo do processo de seleção. Pode-se ainda, dada a avaliação da comissão de seleção, determinar tais prioridades entre os pares de critérios, sendo esta abordagem vinculada ao edital de convocação pública como procedimento.

A Tabela 3 apresenta as escalas adotadas entre os critérios para aplicação do método AHP. Neste estudo de caso, foram arbitradas tais prioridades na tentativa justificada de manter a coerência e equidade nos procedimentos.

Tabela 3

Escalas adotadas entre os critérios para aplicação do método AHP

| Critério 1 | Critério 2 | Escala adotada |
|-----------------------------------|-----------------------------------|-----------------------|
| Experiência (CI) e Escrita (CI) | Objetivos (CI) e Escrita (CI) | 3 |
| Experiência (CI) e Escrita (CI) | Escrita (CI) e Qualificações (CL) | 2 |
| Experiência (CI) e Escrita (CI) | Escrita (CI) e Escrita (CL) | 1 |
| Experiência (CI) e Escrita (CI) | Objetivos (CI) e Escrita (CL) | 1 |
| Objetivos (CI) e Escrita (CI) | Escrita (CI) e Qualificações (CL) | 1 |
| Objetivos (CI) e Escrita (CI) | Escrita (CI) e Escrita (CL) | 1 |
| Objetivos (CI) e Escrita (CI) | Objetivos (CI) e Escrita (CL) | 1 |
| Escrita (CI) e Qualificações (CL) | Escrita (CI) e Escrita (CL) | 1 |
| Escrita (CI) e Qualificações (CL) | Escrita (CI) e Escrita (CL) | 5 |
| Escrita (CI) e Qualificações (CL) | Objetivos (CI) e Escrita (CL) | 4 |

Nota: CI = Carta de Intenção; CL = Currículo Lattes.

A experiência profissional previamente documentada é considerada mais crítica para o sucesso do que as intenções futuras. Isso pode justificar um ajuste para favorecer a experiência sobre os objetivos, ambos descritos na carta de intenção. Por isso, atribuiu-se um valor de 3 na escala AHP, indicando que a experiência e a qualidade da escrita são moderadamente mais importantes.

A qualidade do currículo *Lattes* é elemento importante ao processo, mas a experiência específica articulada na carta é percebida como um melhor indicador de desempenho futuro. Nesse caso, pode-se dar um pouco mais de peso à experiência. Portanto, deu-se um valor de 2, mostrando uma importância ligeiramente maior. Qualificações documentadas no currículo *Lattes* podem ser vistas como uma base sólida para a tomada de decisão, indicando conquistas passadas e habilidades verificáveis. Logo, atribui-se um valor de 5, indicando uma importância fortemente maior.

A consistência na qualidade da escrita entre diferentes documentos é vista como um indicador de atenção e rigor, isso pode ser considerado mais importante do que os objetivos expressos. Indica-se aqui um valor de 4, refletindo uma importância maior. Já considerando que a consistência na qualidade da escrita pode ser um indicador de comunicação eficaz, importante para o ambiente acadêmico, atribui um valor de 1, indicando igual importância, assumindo que ambos são importantes, mas não um mais do que o outro.

A conclusão do processo resultou na classificação final dos candidatos ao apresentar a relação dos candidatos e suas respectivas pontuações obtidas em ordem decrescente, sendo os primeiros colocados com a melhor pontuação no pleito.

4 APRESENTAÇÃO DOS RESULTADOS

Nesta seção, apresentamos os resultados da aplicação do método AHP-TOPSIS-2N na classificação dos candidatos em um processo seletivo competitivo. Foram analisados 62 candidatos, dos quais destacamos as 15 primeiras colocações com base em critérios cuidadosamente selecionados e ponderados. A Tabela 4 mostra os resultados da primeira normalização, enquanto a Tabela 5 apresenta os resultados do segundo procedimento de normalização.

Na Tabela 4, observa-se que o candidato 2 destaca-se com uma pontuação relativa de satisfação (RS) de 0.9810, seguido pelo candidato 5 (RS de 0.5629) e 36 (RS de 0.5286). Estes resultados refletem uma combinação de proximidade à solução ideal positiva e distância da solução ideal negativa, com o 2 apresentando uma distância D^+ significativamente baixa (0.0018) e uma distância D^- relativamente alta (0.0904), sugerindo um perfil bem alinhado aos critérios estabelecidos.

Tabela 4

Avaliação de candidatos em processo seletivo a partir do procedimento 1

| ID Candidato | D+ | D- | RS |
|--------------|--------|--------|--------|
| 2 | 0.0018 | 0.0904 | 0.9810 |
| 5 | 0.0432 | 0.0556 | 0.5629 |
| 36 | 0.0433 | 0.0485 | 0.5286 |
| 41 | 0.0588 | 0.0337 | 0.3642 |
| 11 | 0.0655 | 0.0339 | 0.3410 |
| 3 | 0.0629 | 0.0297 | 0.3210 |
| 40 | 0.0670 | 0.0273 | 0.2897 |
| 23 | 0.0686 | 0.0243 | 0.2619 |
| 48 | 0.0798 | 0.0281 | 0.2604 |
| 37 | 0.0694 | 0.0236 | 0.2540 |
| 33 | 0.0714 | 0.0218 | 0.2334 |
| 10 | 0.0715 | 0.0216 | 0.2323 |
| 31 | 0.0721 | 0.0211 | 0.2263 |
| 55 | 0.0748 | 0.0195 | 0.2064 |
| 42 | 0.0843 | 0.0205 | 0.1954 |

Nota: D+ = distância para a solução ideal positiva; D- = distância para a solução ideal negativa; RS = proximidade relativa.

A Tabela 5 reforça a superioridade do candidato 2 com uma RS de 0.9790, mesmo sob um regime de normalização alternativo, o que indica a robustez dos resultados. O candidato 5 mantém-se em posição proeminente (RS de 0.6223), seguido pelo candidato 36 (RS de 0.5821). A consistência na classificação dos candidatos entre as duas normalizações sugere que a metodologia AHP-TOPSIS-2N é resiliente a variações na abordagem de normalização dos dados.

Tabela 5

Avaliação de candidatos em processo seletivo a partir do procedimento 1

| ID Candidato | D+ | D- | RS |
|--------------|--------|--------|--------|
| 2 | 0.0099 | 0.4616 | 0.9790 |
| 5 | 0.1949 | 0.3210 | 0.6223 |
| 36 | 0.2008 | 0.2797 | 0.5821 |
| 41 | 0.2764 | 0.2158 | 0.4384 |
| 11 | 0.3057 | 0.2222 | 0.4209 |
| 3 | 0.2971 | 0.1992 | 0.4013 |
| 40 | 0.3324 | 0.1840 | 0.3564 |
| 23 | 0.3264 | 0.1775 | 0.3523 |
| 37 | 0.3304 | 0.1747 | 0.3459 |
| 48 | 0.3807 | 0.1975 | 0.3416 |
| 33 | 0.3410 | 0.1676 | 0.3296 |
| 10 | 0.3416 | 0.1672 | 0.3287 |
| 31 | 0.3447 | 0.1652 | 0.3239 |
| 32 | 0.3649 | 0.1530 | 0.2954 |
| 17 | 0.3697 | 0.1503 | 0.2891 |

Nota: D+ = distância para a solução ideal positiva; D- = distância para a solução ideal negativa; RS = proximidade relativa.

É importante notar que, apesar da variação nos valores absolutos das distâncias D+ e D- entre os dois procedimentos, a ordem relativa dos candidatos permaneceu quase inalterada, indicando a confiabilidade do método

em determinar a classificação dos candidatos. Os candidatos que aparecem nas primeiras posições demonstraram um perfil holístico que equilibra tanto a excelência documentada em seu currículo *Lattes* quanto a articulação e clareza de objetivos e experiências em suas cartas de apresentação.

Os resultados obtidos são um testemunho da eficácia do método AHP-TOPSIS-2N em discernir as sutilezas entre os candidatos, proporcionando um julgamento quantitativo que é essencial para uma seleção justa e objetiva. A aplicação deste método revela-se uma ferramenta valiosa para comitês de seleção, permitindo uma comparação multifacetada e uma classificação discriminante entre um grande número de candidatos.

5 DISCUSSÕES

5.1 Quanto aos aspectos metodológicos

Os métodos de normalização utilizados no AHP-TOPSIS-2N são utilizados para padronizar os valores dos critérios, permitindo que eles sejam comparados em uma escala comum, independentemente de suas unidades originais ou magnitudes. Normalmente, existem dois métodos comuns de normalização: o linear e o vetorial. No método linear, os valores são ajustados para uma escala comum, frequentemente de 0 a 1, onde 0 representa a menor preferência e 1 a maior. No método do vetor, os valores são divididos pelo vetor norma dos critérios, resultando em uma escala unitária que preserva as relações de proporcionalidade.

A discussão dos resultados obtidos através do método AHP-TOPSIS-2N deve considerar a relevância e o peso de cada estratégia prevista no processo seletivo. Por exemplo, se critérios como experiência profissional e qualificações acadêmicas são altamente ponderados, a classificação dos candidatos deve refletir a prioridade dada a estes aspectos. A consistência dos resultados entre diferentes métodos de normalização indica que os candidatos selecionados como os mais adequados possuem atributos que são robustos e consistentes, independentemente do método de normalização aplicado.

A interpretação deste contexto no processo seletivo deve considerar que os resultados obtidos pelos candidatos refletem uma agregação quantitativa das avaliações de múltiplos critérios, traduzindo-se em uma medida de sua justiça global para a carga ou função em questão. Candidatos com resultados mais altos são considerados mais alinhados com o perfil desejado. Esses candidatos não demonstram apenas excelência individual em determinados critérios, mas também um equilíbrio entre todos os aspectos avaliados. Isso sugere que esses candidatos têm o potencial de desempenhar suas funções de forma mais eficaz e contribuir significativamente para os objetivos da organização ou instituição educacional.

5.2 Quanto à literatura

Na discussão dos achados deste estudo no contexto da literatura existente, os modelos preditivos são utilizados para identificar padrões de desempenho dos alunos e fornecer alertas tanto para os estudantes quanto para os docentes (Funchal et al., 2016; Souza, 2021; Morais et al., 2023; Oliveira & Medeiros, 2024). Nesse cenário, é possível expandir as aplicações de modelos

preditivos como instrumentos para a melhoria dos processos de seleção acadêmica.

Os resultados obtidos através dos métodos alinham-se com os princípios de eficiência, precisão e equidade. A metodologia empregada oferece uma análise rápida das candidaturas, atendendo à demanda por eficiência. Isso se reflete no tempo necessário para analisar e classificar um grande volume de candidatos, uma métrica indicada para avaliar a eficiência do modelo (Diverio et al., 1995). A precisão do AHP-TOPSIS-2N é demonstrada pela sua capacidade de diferenciar sutilezas entre os candidatos, o que está em consonância com as observações de Kuncel et al. (2005) ao enfatizarem a importância de identificação de candidatos verdadeiramente qualificados.

Quanto à equidade, o método AHP-TOPSIS-2N pode ser considerado um avanço no sentido de atender às preocupações de Barreto et al. (2023), já que promove uma avaliação imparcial dos candidatos, baseando-se nas considerações em méritos quantitativos e qualitativos, sem espaço para discriminação subjetiva ou inconsciente. Este aspecto é crítico, visto que a análise de grupos desfavorecidos e a prevenção de vieses são centrais para a equidade do processo (Rosa & Gonçalves, 2015).

No entanto, Barocas e Selbst (2016) alertam que a objetividade de tais modelos está intrinsecamente ligada à qualidade e à imparcialidade dos dados e critérios utilizados. Portanto, os resultados deste estudo também reforçam a necessidade de uma seleção cuidadosa de critérios e uma revisão constante dos dados utilizados. Por isso, esses resultados estão em conformidade com as recomendações de Corbett-Davies et al. (2017) para o desenvolvimento de modelos que promovam justiça e equidade. Ao implementar o método proposto, esta pesquisa também contribui para a literatura que explora formas de controlar processos complexos, uma vez que a metodologia oferece uma abordagem estruturada e transparente que pode ser auditada e ajustada para melhorar continuamente a justiça da seleção.

6 CONSIDERAÇÕES FINAIS

Este estudo abordou o processo de seleção acadêmica por meio do método AHP-TOPSIS-2N, uma abordagem analítica que visa melhorar a eficiência, precisão e equidade da seleção de candidatos em pós-graduações. A literatura atual aponta para uma crescente necessidade de métodos que reduzam o desperdício de tempo e recursos e promovam decisões justas e imparciais. No entanto, uma limitação recorrente nos estudos existentes é a tendência de subestimar as visões inerentes aos dados e aos critérios utilizados para modelos treinados matemáticos de seleção, conforme planejado por Barocas e Selbst (2016).

As contribuições deste artigo residem na aplicação prática de uma metodologia robusta que fornece uma análise quantitativa detalhada dos candidatos, conduzindo um processo de seleção mais fundamentado e transparente. Também foi demonstrado que o método AHP-TOPSIS-2N pode efetivamente classificar um grande número de candidatos com consistência, mesmo quando solicitado a diferentes métodos de normalização, destacando a confiabilidade e a validade da abordagem.

Apesar dos avanços, este estudo não explora o impacto de dados históricos vistos na modelagem final e na decisão. Este é um campo fértil para investigações futuras, pois o reconhecimento e a mitigação de vieses são essenciais para a integridade de qualquer processo seletivo. Além disso, embora se tenha abordado a eficiência e a precisão, a mensuração da equidade

em processos de seleção permanece como uma lacuna de pesquisa, como também uma automatização total, desde a estruturação, tratamento dos dados e aplicação metodológica para os resultados finais. Nessa perspectiva, trabalhos futuros poderiam se concentrar em estudar análises mais refinadas para avaliar e garantir a equidade, particularmente em relação a grupos desfavorecidos.

Como perspectiva de trabalho futuro, sugere-se a implementação de técnicas de aprendizado de máquina para avaliar e corrigir possíveis vieses nos critérios de seleção. Além disso, recomenda-se uma expansão do estudo para incluir comparações interculturais e interinstitucionais para validar a generalização do método AHP-TOPSIS-2N em diferentes contextos acadêmicos. Outra possível frente futura de pesquisa se concentra em um mapeamento de todo o processo para uma proposta de automatização computacional completa. A longo prazo, espera-se que tais melhorias nos modelos de seleção acadêmica contribuam para processos mais céleres, justos e equitativos, alinhados com as necessidades e valores de uma sociedade em constante evolução.

REFERÊNCIAS

- Aguiar, E. S., Silva, V. B. S., & Schramm, F. (2015). Proposta de um procedimento para a seleção de candidatos à bolsa auxílio do PNAES. In *Anais do XLVII - Simpósio Brasileiro de Pesquisa Operacional*, Porto de Galinhas, PE. <https://www.din.uem.br/sbpo/sbpo2015/pdf/142674.pdf>
- Awasthi, A., & Chauhan, S. S. (2011). Using AHP and Dempster-Shafer theory for evaluating sustainable transport solutions. *Environmental Modelling & Software*, 26(6), 787-796.
<https://doi.org/10.1016/j.envsoft.2010.11.010>
- Barreto, J. C., Stefano, E., Freitag, A. E. B., Santana, W. A., & Picoli, M. A. (2023). Currículo inclusivo: uma proposta para equidade e efetividade nos processos de recrutamento. *Revista de Gestão e Secretariado*, 14(5), 7586-7607.
<https://doi.org/10.7769/gesec.v14i5.2139>
- Barocas, S., & Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, 104(3), 671-732. <https://dx.doi.org/10.2139/ssrn.2477899>
- Behzadian, M., Otaghsara, S. K., Yazdani, M., & Ignatius, J. (2012). A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications*, 39(17), 13051-13069.
<https://doi.org/10.1016/j.eswa.2012.05.056>
- Boettiger, C. (2015). An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*, 49(1), 71-79.
<https://doi.org/10.1145/2723872.2723882>
- Bozza, G., Ruy, M. M., Santos, M., Moreira, M. A. L., Rocha Junior, C. S., & Gomes, C. F. S. (2020). *Three Decision Methods (3DM) Software Web* (v. 1). <https://www.3decisionmethods.com/3DM/index.html>
- Brynjolfsson, E., & Hitt, L. M. (2000). Beyond Computation: Information Technology, Organizational Transformation and Business Performance. *Journal of Economic Perspectives*, 14(4), 23-48.
<https://doi.org/10.1257/jep.14.4.23>
- Burak, D. (1992). *Modelagem Matemática: ações e interações no processo de ensino-aprendizagem* [Tese de Doutorado, Universidade Estadual de Campinas]. Repositório Institucional da UNICAMP.
<https://www.psiem.fe.unicamp.br/content/modelagem-matematica-acoes-e-interacoes-no-processo-de-ensino-aprendizagem>
- Cassiano, C. N., Lima, L. C., & Zuppani, T. S. (2016). A eficiência das redes sociais em processos de recrutamento organizacional. *NAVUS - Revista de Gestão e Tecnologia*, 6(2), 52-67.
<https://doi.org/10.22279/navus.2016.v6n2.p52-67.362>
- Chiavenato, I. (2009). *Planejamento, recrutamento e seleção de pessoal: como agregar talentos à empresa*. (7. ed.). Manole.
- Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic Decision Making and the Cost of Fairness. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY. <https://doi.org/10.1145/3097983.3098095>
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. (2010). *Relatório de Avaliação 2007-2009, Trienal 2010*.
http://trienal.capes.gov.br/wp-content/uploads/2011/02/RELAT%C3%93RIO-DE-AVALIA%C3%87%C3%83O_ADMINISTRA%C3%87%C3%83O.pdf
- Cunha, L. A. C. R. (1974). A pós-graduação no Brasil: função técnica e função social. *Revista de Administração de Empresas*, 14(5), 66-70.
<https://doi.org/10.1590/S0034-75901974000500006>
- Diverio, T., Navaux, P., & Claudio, D. (1995). Alto Desempenho e Eficiência em Processamento Numérico. *Anais do VII Simpósio de Arquitetura de Computadores e Processamento de Alto Desempenho*, Porto Alegre, RS.
<https://doi.org/10.5753/sbac-pad.1995.19867>
- Domingues, J., & Domingues, D. M. (2023). Desvendando o passado: autenticidade de documentos digitalizados na realização de uma narrativa histórica na educação matemática. *Convergências: Estudos em*

- Humanidades Digitais*, 1(3), 66-80.
<https://doi.org/10.59616/cehd.v1i03.617>
- Fernandes, D. Y. S., & Rêgo, A. S. C. (2023). Viés, ética e responsabilidade social em modelos preditivos. *Computação Brasil*, (51), 19-23. <https://doi.org/10.5753/compbr.2023.51.3988>
- Figueiredo Filho, D. B., Rocha, E. C., Paranhos, R., Silva, A. H., Silva Junior, J. A., Oliveira, L. E., & Alves, D. P. (2014). Análise fatorial garantida ou o seu dinheiro de volta: uma introdução à redução de dados. *Revista Eletrônica de Ciência Política*, 5(2), 185-211. <http://dx.doi.org/10.5380/recp.v5i2.40368>
- Funchal, J. P. S., Rodrigues, A. S. P., & Borges, E. N. (2016). Um modelo preditivo para estudo da evasão na graduação utilizando mineração de dados. *RETEC - Revista de Tecnologias*, 9(3), 74-84.
<https://www.fatecourinhos.edu.br/retec/index.php/retec/article/view/262>
- Kahneman, D. (2012). *Rápido e devagar: duas formas de pensar*. (1. ed.). Objetiva.
- Gil, A. C. (2010). *Como elaborar projetos de pesquisa*. (5. ed.). Atlas.
- Hwang, CL., & Yoon, K. (1981). Methods for Multiple Attribute Decision Making. In *Multiple Attribute Decision Making*. Springer.
- Igarashi, D. C. C., Ensslin, S. R., Ensslin, L., & Paladini, E. P. (2008). A qualidade do ensino sob o viés da avaliação de um programa de pós-graduação em contabilidade: proposta de estruturação de um modelo híbrido. *Revista de Administração*, 43(2), 117-137.
<https://www.revistas.usp.br/rausp/article/view/44471>
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The Validity of Self-Reported Grade Point Averages, Class Ranks, and Test Scores: A Meta-Analysis and Review of the Literature. *Review of Educational Research*, 75(1), 63-82. <https://doi.org/10.3102/00346543075001063>
- Lei n. 9.394, de 20 de dezembro de 1996. (1996). Estabelece as diretrizes e bases da educação nacional. Brasília, DF.
- Lei n. 13.709, de 14 de agosto de 2018. (2018). Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF.
- Leite, D., Tutikian, J., & Holz, N. (2000). *Avaliação e compromisso: construção e prática da avaliação institucional em uma universidade pública*. Editora da Universidade Federal do Rio Grande do Sul.
- Morais, D. M. G., Souza, H. R., & Silva, V. B. (2023). Modelo preditivo do desempenho acadêmico: um estudo para a disciplina de Cálculo Diferencial e Integral para Engenharias. *FTT Journal of Engineering and Business*, 1(8), 47-68.
<https://saijournal.cefsa.org.br/index.php/FTT/article/view/509>
- Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2014). *Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining*. John Wiley & Sons.
- Neeraj, Goraya, M. S., & Singh, D. (2019). A Ranking Based Model for Selecting Optimum Cloud Geographical Region. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(10), 793-797. <https://doi.org/10.35940/ijitee.J8908.0881019>
- Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175-220.
<https://doi.org/10.1037/1089-2680.2.2.175>
- Nobre, D. K. (2020). Processo Penal e Direito Humanos: notas iniciais sobre sistemas que utilizam inteligência artificial em julgamento. In De Souza Junior, A. B., Ferreira, D. D., Almeida, M. P., & Filpo, K. P. L.(Orgs.), *Dimensões jurídicas dos direitos humanos - vol. 4* (pp. 83-92). Pembroke Collins.
- Oliveira, R. S., & Medeiros, F. P. A. (2024). Modelo de Predição de Evasão Escolar com Base em Dados de Autoavaliação de Cursos de Graduação. *Revista Brasileira de Informática na Educação*, 32, 1-21.
<https://doi.org/10.5753/rbie.2024.3542>

- Oliveira, R., Montez, C., & Wangham, M. (2020). Projeto Integrado de Controle de Pelotões baseado em Consenso e um Protocolo de Disseminação de Dados Confiável. *Anais do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, Porto Alegre, RS. <https://doi.org/10.5753/sbrc.2020.12302>
- Rocha, C. J., Porto, L. V., & Abaurre, H. E. (2020). Discriminação algorítmica no trabalho digital. *Revista de Direitos Humanos e Desenvolvimento Social*, 1, 1-21. <https://doi.org/10.24220/2675-9160v1e2020a5201>
- Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384. <https://doi.org/10.1016/j.compedu.2007.05.016>
- Rosa, C. M., & Gonçalves, A. M. (2015). A política de cotas na UFG: desvelando o perfil dos estudantes cotistas. *Revista NUPEM*, 7(1), 47-66. <https://doi.org/10.33871/nupem.v7i1.251>
- Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill.
- Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), 83-98. <https://doi.org/10.1504/IJSSCI.2008.017590>
- Santos, D. J., & Machado, M. A. S. (2018). A multicriterary approach to decision-making in personal allocation using the todo method. *International Contemporary Management Review*, 1(1), 1-19. <https://www.icmreview.com/icmr/article/view/17>
- Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 72, 414-422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Silva, D. J. M., & Silva, M. A. (2016). Perfil dos processos seletivos dos programas de pós-graduação stricto sensu em ciências contábeis no brasil. *Revista CAMINE: Caminhos da Educação*, 8(1), 88-118. <https://periodicos.franca.unesp.br/index.php/caminhos/article/view/1709>
- Soares, S. V., Will, A. R., Miranda, M. P., & Pfitscher, E. D. (2011). Pós-graduação em ciências contábeis no brasil: contexto e processo de seleção. *Revista Ambiente Contábil*, 3(2), 59-74. <https://periodicos.ufrn.br/ambiente/article/view/1357>
- Souza, R. R., Lima, M. V. A., & Correa, A. C. (2014). Influência da concessão de bolsa de estudos na produtividade acadêmica dos estudantes de administração ao nível pós-graduação stricto sensu no brasil. *XIV Colóquio Internacional de Gestão Universitária*, Florianópolis, SC. <https://core.ac.uk/download/pdf/30408105.pdf>
- Souza, V. F. (2021). Mineração de dados educacionais com aprendizagem de máquina. *Revista Educar Mais*, 5(4), 766-787. <https://doi.org/10.15536/reducarmais.5.2021.2417>
- Strauss, E., Bôas Júnior, M. V., & Ferreira, W. L. L. (2022). A importância de utilizar métricas adequadas de avaliação de performance em modelos preditivos de machine learning. *Projectus*, 7(2), 52-62. <https://doi.org/10.15202/25254146.2022v7n2p52>
- Takakura, F. I., & Duarte, L. G. M. (2022). Inteligência artificial no direito: dilemas e contribuições. *Revista de Direito, Inovação, Propriedade Intelectual e Concorrência*, 8(1), 1-23. <https://www.indexlaw.org/index.php/revistadipic/article/view/8633>
- Thorndike, E. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4(1), 25-29. <http://dx.doi.org/10.1037/h0071663>
- Trevizano, W. A., & Freitas, A. L. P. (2005). Emprego do Método da Análise Hierárquica (A.H.P.) na seleção de processadores. *XXV Encontro Nacional de Engenharia de Produção*, Porto Alegre, RS. <https://doi.org/10.13140/2.1.3576.3201>

Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A comparative Study*. (1. ed.). Springer.

Zeleny, M. (1982). *Multiple Criteria Decision Making*. McGraw-Hill.