em foco
Revista Científica em Educação a Distância

Análise de Perfis Metacognitivos para Redução da Evasão em EAD: uma Abordagem Baseada em Ciência de Dados

Analysis of Metacognitive Profiles to Reduce Dropout in Distance Learning: a Data Science-Based Approach

Veronica Gabriela Ribeiro DA SILVA[1]*
Arthur Gonsales da SILVA[2]
Caio de Oliveira BASTOS[1]
Bruno Duarte GOMES[1]

[1]Universidade Federal do Pará. Rua Augusto Corrêa, 01 – Belém – PA – Brasil.

[2]Universidade de Campinas. Rua Vital Brasil, 888 – Campinas – SP – Brasil.

Resumo. Este estudo investiga a integração da análise metacognitiva em neuroeducação e ciência de dados para aprimorar a experiência da educação a distância (EAD). Foram analisados dados de 1391 estudantes de pós-graduação da Faculdade Estratego, utilizando um questionário abrangente e o Inventário de Consciência Metacognitiva (MAI) para avaliar dimensões acadêmicas, pessoais, socioeconômicas e metacognitivas. O pré-processamento e análise dos dados foram realizados utilizando Python no Jupyter Notebook, empregando o agrupamento K-Means e a classificação por Máquina de Vetores de Suporte (SVM) para identificar e categorizar perfis metacognitivos. Nossos resultados revelam quatro perfis metacognitivos distintos entre os estudantes, que se correlacionam significativamente com seu desempenho acadêmico e engajamento na EAD. O processo de agrupamento, validado pelos métodos Elbow e Análise de Silhueta, confirmou a formação ideal desses perfis, aprimorando nossa compreensão das diversas estratégias de aprendizagem dentro da população estudantil. A Análise de Componentes Principais (PCA) foi utilizada para refinar ainda mais os dados, focando nos atributos metacognitivos mais significativos que influenciam os resultados de aprendizagem. Esta análise destacou estratégias metacognitivas específicas fundamentais para promover ambientes de aprendizagem eficazes e reduzir as taxas de abandono. O estudo ressalta o potencial de

PRELO

combinar princípios de neuroeducação, insights metacognitivos e técnicas de ciência de dados para personalizar estratégias educacionais que acomodem perfis de aprendizes diversos. Alinhando o conteúdo educacional e as metodologias às capacidades cognitivas e metacognitivas dos estudantes, os educadores podem melhorar significativamente o engajamento e a eficácia nos ambientes de EAD, melhorando, em última análise, o sucesso acadêmico e as taxas de retenção.

Palavras-chave: Neuroeducação. Metacognição. Ciência de dados. Educação a distância. Perfis cognitivos.

Abstract. *This study investigates the integration of neuroeducation, metacognitive analysis, and data science to enhance the digital education (DE) experience. We collected and analyzed data from 1,391 postgraduate students at Estratego College, utilizing a comprehensive questionnaire and the Metacognitive Awareness Inventory (MAI) to assess academic, personal, socioeconomic, and metacognitive dimensions. Data preprocessing and analysis were conducted using Python in Jupyter Notebook, employing K-Means clustering and Support Vector Machine (SVM) classification to identify and categorize metacognitive profiles. Our findings reveal four distinct metacognitive profiles among students, which correlate significantly with their academic performance and engagement in DE. The clustering process, validated through the Elbow Method and Silhouette Analysis, confirmed the optimal formation of these profiles, enhancing our understanding of the diverse learning strategies within the student population. Principal Component Analysis (PCA) was utilized to further refine the data, focusing on the most significant metacognitive attributes that influence learning outcomes. This analysis highlighted specific metacognitive strategies that are pivotal in fostering effective learning environments and reducing dropout rates. The study underscores the potential of combining neuroeducation principles, metacognitive insights, and data science techniques to tailor educational strategies that accommodate diverse learner profiles. By aligning educational content and methodologies with students' cognitive and metacognitive capacities, educators can significantly enhance engagement and efficacy in DE settings, ultimately improving academic success and retention rates. This research contributes to the field by demonstrating how data-driven insights can inform and transform educational practices, offering a model for leveraging technology and cognitive science to meet the evolving demands of digital education.*

Keywords: *Neuroeducation. Metacognition. Data science. Digital education. Cognitive profiles.*

Recebido: 24 /09/2024 Aceito: 06/06/2025 Publicado: 02/07/2025

Editores Responsáveis: Daniel Salvador/ Carmelita Portela

1. Introdução

1.1. Neuroeducação e Análise Metacognitiva na Educação a Distância (EAD)

Nos últimos anos, os avanços científicos e tecnológicos abriram novos horizontes na compreensão e gestão de dados e comportamentos. A neurociência, em particular, emergiu como uma disciplina-chave, desvendando mistérios fundamentais do cérebro humano e impulsionando o progresso em várias áreas. Simultaneamente, a Educação a Distância (EAD) tem se tornado cada vez mais significativa na educação contemporânea, oferecendo flexibilidade e acessibilidade que atraem um número crescente de estudantes em programas de treinamento profissional, ensino superior e educação continuada. No entanto, a EAD também apresenta desafios únicos, como a manutenção de altos níveis de engajamento e a prevenção de taxas de abandono (Orsucci; Sala, 2022).

A neuroeducação, que integra princípios da neurociência à prática educacional, fornece *insights* valiosos para otimizar a experiência de aprendizagem (Cosenza; Guerra, 2011). No contexto da EAD, entender funções cerebrais — como processos de memória, atenção e motivação — pode tornar o design de cursos online mais eficazes (Cosenza; Guerra, 2011; Filatro; Cairo, 2015; Orsucci; Sala, 2022).

Uma área particularmente promissora dentro da neuroeducação é a análise metacognitiva. Metacognição refere-se à capacidade de um indivíduo monitorar e regular seus processos de pensamento (Flavell, 1979). Segundo Alliprandini *et al.* (2014), o sucesso na educação a distância depende da capacidade de aprender a aprender. Isso envolve ajudar os estudantes a desenvolver uma compreensão mais profunda de seus processos de aprendizagem e a melhorar suas habilidades de estudo (Alliprandini *et al*., 2014; Santos *et al*., 2021).

As teorias educacionais modernas e a própria EAD promovem qualidades como proatividade, autonomia e autoconsciência. Para otimizar sua aprendizagem, os estudantes precisam entender seus objetivos de aprendizagem e atividades educacionais, permitindo-lhes criar e reconhecer diferentes estratégias para adquirir, analisar e regular seu aprendizado (Flavell, 1979). O desempenho eficaz na EAD requer que os estudantes sejam participantes ativos, empregando pensamento crítico, disposição para buscar conhecimento e responsabilidade na aplicação e controle dessas habilidades. Portanto, é essencial avaliar as estratégias cognitivas e metacognitivas que influenciam a relação entre os estudantes, os recursos educacionais e seus ambientes virtuais de aprendizagem (Rivas *et al*., 2022; Santos *et al.*, 2021).

Abbad, Carvalho e Zerbini (2004) pesquisaram o perfil dos estudantes geralmente envolvidos na EAD, alinhando suas expectativas com os desafios enfrentados por atores sociais em diversos contextos educacionais. Baseado na literatura produzida por Filatro (2004, 2008 e 2021), é possível inferir e articular várias maneiras pelas quais a neuroeducação e a análise metacognitiva podem influenciar positivamente a experiência de aprendizagem na EAD:

- **Design Instrucional Baseado em Evidências:** Entender os fundamentos neurocientíficos da aprendizagem otimiza o design de cursos online, incluindo a criação de materiais de aprendizagem visualmente atraentes que estimulem a atenção e a integração de elementos interativos para promover o engajamento.

- **Personalização da Aprendizagem:** A análise metacognitiva permite que os estudantes identifiquem suas estratégias de aprendizagem mais eficazes, possibilitando que os cursos online sejam adaptados às preferências de aprendizagem individuais, tornando a experiência mais relevante e envolvente.

- **Apoio à Tomada de Decisão:** Através da metacognição, os estudantes aprendem a avaliar seu progresso e identificar áreas que precisam de mais suporte, ajudando a prevenir taxas de abandono ao enfrentar desafios antes que se tornem insuperáveis.

- ***Feedback* Baseado em Dados:** A análise metacognitiva gera dados valiosos sobre o desempenho e processos de pensamento dos estudantes, permitindo que instrutores e designers de cursos forneçam feedback mais preciso e direcionado, facilitando a melhoria contínua.

- **Fomento à Autonomia:** A metacognição capacita os estudantes a se tornarem aprendizes mais autônomos, desenvolvendo a habilidade de autorregular seu aprendizado, uma habilidade valiosa para a aprendizagem ao longo da vida.

Um grande dilema educacional é como o processo metacognitivo influencia tanto a experiência de aprendizagem em ambientes virtuais quanto as taxas de abandono na EAD. O último censo da Associação Brasileira de Educação a Distância (ABED), publicado em abril de 2022, destacou que as taxas de abandono são multifacetadas e foram exacerbadas pela pandemia de COVID-19. Os dados mostraram uma taxa de abandono de 25% em cursos de graduação, pós-graduação e livres, principalmente devido a dificuldades financeiras e à incapacidade de se adaptar a novas metodologias. Estudos internacionais recentes também apontam que, além de fatores econômicos e estruturais, elementos cognitivos como a autorregulação e a consciência metacognitiva são determinantes no engajamento e na persistência de estudantes em contextos de aprendizagem online (Pelikan *et al.*, 2021).

Apesar dos programas de controle de abandono, um problema significativo permanece a falta de dados suficientes para a tomada de decisões e consolidação de estratégias por gestores,

professores e estudantes. Abordar esse desafio requer o desenvolvimento e a disponibilização de estratégias e recursos para grupos com necessidades variadas. Isso exige um conhecimento aprofundado do público através da coleta e análise de dados, identificando características específicas. A Análise Contextual (AC) torna-se crucial no processo de design educacional, avaliando os recursos disponíveis, as limitações, estratégias e custo-benefício (Filatro, 2015).

A AC também apoia a avaliação metacognitiva ao oferecer aos estudantes informações relevantes sobre seu desempenho educacional, permitindo-lhes interpretar, entender e tomar decisões para otimizar seu processo de aprendizagem (Gutierrez-Pachas *et al.*, 2023; Mory, 2004).

1.2. Ciência de Dados Aplicada à Análise Contextual e Metacognitiva na Educação a Distância

As Ciências da Informação, particularmente a Ciência de Dados, oferecem um potencial significativo para avanços educacionais ao analisar e processar grandes quantidades de dados para obter *insights* sobre fenômenos sociais e técnicos (Filatro, 2021). A Análise de Aprendizagem, um ramo da Ciência de Dados, permite a análise quantitativa e qualitativa das atividades educacionais, fornecendo aos educadores e gestores ferramentas para visualizar informações e melhorar a tomada de decisões e ajustes de estratégias (Filatro, 2021). No entanto, como observado por Romero e Ventura (2013), a coleta e gestão de dados continuam sendo desafios significativos para as instituições educacionais.

Para abordar esse dilema, Rabelo *et al.* (2017) discutem como a técnica de Mineração de Dados Educacionais (MDE) permite a criação e utilização de modelos para visualizar padrões e gerar *insights* a partir de um conjunto de dados resultantes dos processos de ensino-aprendizagem. Além de seu valor científico, essa técnica também possui impacto social, pois contribui para decisões mais eficientes em políticas educacionais e práticas pedagógicas (Zhang *et al.*, 2024). A integração de abordagens baseadas em Análise Contextual, aliada ao conhecimento proveniente das Neurociências e ao uso de tecnologias digitais, pode aprofundar a compreensão dos processos de ensino-aprendizagem — especialmente em ambientes virtuais mediados por Tecnologias da Informação e Comunicação (TICs).

Além disso, a combinação de estratégias de Análise Contextual com o conhecimento Neurocientífico e técnicas de Ciência de Dados pode identificar padrões e conexões nos processos de ensino-aprendizagem, especialmente em contextos de educação a distância mediados por tecnologias de informação e comunicação (TICs). Esta abordagem pode fomentar melhores práticas, otimizar resultados e melhorar a tomada de decisões no campo educacional (Silva Júnior *et al.*, 2024).

Através da análise contextual (AC), por exemplo, alguns pontos são essenciais para qualquer ação educacional, como:

- Localização demográfica, perfil cultural e outras características inerentes;
- Trajetória de carreira e experiências;
- Conhecimento prévio dos tópicos do curso;
- Habilidades e competências para o uso da Internet e outros recursos digitais;
- Características cognitivas e comportamentais: hábitos de estudo, estratégias e estilos de aprendizagem;
- Características motivacionais: motivações pessoais, acadêmicas e profissionais e o valor que tal processo educacional inserido em seu público pode trazer para eles.

Este trabalho tem como objetivo principal integrar análise metacognitiva, neuroeducação e ciência de dados para otimizar experiências de aprendizagem na Educação a Distância. Especificamente, busca-se identificar e categorizar perfis metacognitivos distintos entre estudantes de pós-graduação através de técnicas avançadas de agrupamento de dados, analisar as características de cada perfil identificado, estabelecer relações entre esses perfis e o comportamento de aprendizagem dos estudantes, e propor estratégias pedagógicas personalizadas fundamentadas em princípios neurocientíficos. A pesquisa visa ainda compreender como esses diferentes perfis metacognitivos podem influenciar as taxas de evasão na EAD, contribuindo para o desenvolvimento de abordagens educacionais mais adaptativas que considerem a diversidade cognitiva dos estudantes.

2. Metodologia

2.1. Coleta de Dados

A pesquisa visou analisar dados acadêmicos, perfis de estudo, dados pessoais, socioeconômicos e metacognitivos de 1391 estudantes de pós-graduação do nível lato sensu (especializações e MBA) em uma instituição parceira, a Faculdade Estratego, sediada em Belém - Pará. Esta instituição oferece cursos nas áreas de educação, gestão, saúde e tecnologia em todo o país na modalidade de Educação a Distância (EAD).

Para a análise contextual, foram utilizados dados obtidos das experiências educacionais dos estudantes dentro de seu Ambiente Virtual de Aprendizagem (AVA), denominado Sala de Aula Virtual (SAV) na Faculdade Estratego. Um questionário composto por 35 perguntas (Pesquisa de Perfil Acadêmico) focando em aspectos acadêmicos, pessoais e socioeconômicos foi

desenvolvido e disponibilizado através do Google Forms (Google LLC) para os estudantes em seu AVA.

O questionário de Pesquisa de Perfil Acadêmico é dividido em três seções:

- **Dados Acadêmicos:** Esta seção refere-se ao histórico acadêmico diretamente relacionado ao curso do estudante e sua formação educacional até a conclusão do questionário.
- **Perfil de Estudo:** Esta seção coleta informações sobre como os estudantes estudam, os dispositivos que utilizam, suas preferências e fatores que influenciam a qualidade da assimilação do conteúdo dentro e fora do AVA.
- **Perfil Pessoal e Socioeconômico:** Esta seção visa coletar informações mais pessoais, como idade, educação, estado civil, renda familiar e outros detalhes relevantes para entender seu contexto e realidade.

Subsequentemente, foram coletados dados dos estudantes que participaram da fase inicial da pesquisa através do Inventário de Consciência Metacognitiva (MAI) — Metacognitive Awareness Inventory (Adaptado de Lima Filho e Bruni, 2015; Schraw e Dennison, 1994) — um questionário de autoavaliação metacognitiva adequado para medir a consciência metacognitiva em adolescentes e adultos. Consiste em 52 perguntas e é bem validado na literatura (Questionário - Inventário de Consciência Metacognitiva (MAI)).

O MAI avaliou dois aspectos cruciais da metacognição:

- **Conhecimento Metacognitivo:** Inclui conhecimento declarativo (conhecimento sobre si mesmo e sobre estratégias), conhecimento processual (conhecimento sobre como executar as estratégias) e conhecimento condicional (conhecimento sobre quando e por que usar as estratégias).
- **Regulação da Cognição:** Envolve planejamento, monitoramento e avaliação das atividades cognitivas.

2.2. Análise de Dados

O pré-processamento e análise dos dados foram conduzidos utilizando Python em um ambiente Jupyter Notebook. As seguintes etapas foram realizadas:

- **Limpeza e Pré-processamento dos Dados:** Isso envolveu o tratamento de valores ausentes, normalização dos dados e codificação de variáveis categóricas.
- **Agrupamento:** O agrupamento K-Means foi empregado para categorizar os estudantes em perfis metacognitivos distintos. O número ótimo de grupos ou

clusters foi determinado utilizando o Método Cotovelo (*Elbow*) e a Análise de Silhueta.

- **Classificação:** A classificação por Máquina de Vetores de Suporte (SVM) foi usada para explorar a relação entre os perfis metacognitivos identificados e os resultados acadêmicos dos estudantes.
- **Redução de Dimensionalidade:** A Análise de Componentes Principais (PCA) foi utilizada para reduzir a dimensionalidade dos dados, focando nos atributos metacognitivos mais significativos que influenciam os resultados de aprendizagem.

O processo de agrupamento revelou quatro perfis metacognitivos distintos entre os estudantes, que foram significativamente correlacionados com seu desempenho acadêmico e engajamento na EAD. A PCA refinou ainda mais os dados, destacando estratégias metacognitivas específicas que são fundamentais para promover ambientes de aprendizagem eficazes e reduzir as taxas de abandono.

2.3. Validação

O processo de agrupamento foi validado através do método cotovelo e da análise de silhueta, confirmando a formação ideal dos perfis metacognitivos. O desempenho do modelo de classificação SVM foi avaliado utilizando métricas como acurácia, precisão, *recall* e *F1-score*.

Integrando princípios de neuroeducação, *insights* metacognitivos e técnicas de ciência de dados, este estudo fornece uma compreensão abrangente das diversas estratégias de aprendizagem dentro da população estudantil, oferecendo um modelo para adaptar estratégias educacionais para aumentar o engajamento e a eficácia em ambientes de EAD.

3. Resultados

3.1. Processo de Clusterização

3.1.1. *K-Means*

O *K-Means* é uma técnica para agrupar dados, buscando identificar grupos onde os elementos dentro de cada grupo são mais semelhantes entre si do que aos de outros grupos. Ao analisar os grupos formados pelo *K-Means*, podemos entender melhor as características e padrões presentes nos dados, facilitando a tomada de decisões e a geração de insights. Essa técnica é amplamente utilizada em diversas áreas, como análise de dados, reconhecimento de padrões e segmentação de clientes, devido à sua eficiência e capacidade de lidar com grandes volumes de

dados. Para garantir a qualidade dos agrupamentos, utilizamos dois métodos de validação: o Método *Elbow* e a Análise de Silhueta.

3.1.2. Método do Cotovelo (*Elbow*)

O método do cotovelo, ou simplesmente método cotovelo, também conhecido pelo seu nome em Inglês *Elbow*, é uma técnica visual bastante utilizada para determinar o número ideal de grupos (*clusters*) ao aplicar o algoritmo *K-Means* em um conjunto de dados. De forma didática, imagine que estamos tentando separar diferentes tipos de frutas em caixas, mas não sabemos ao certo quantas categorias existem. O método do cotovelo nos ajuda a responder essa pergunta. Para isso, testamos diferentes quantidades de grupos e, para cada uma delas, calculamos uma medida chamada soma das distâncias quadradas dentro dos grupos, que indica o quão próximos estão os dados de cada grupo em relação ao seu centro. Em termos práticos, quanto menor essa soma, mais compactos e bem definidos estão os grupos.

Ao representar esses resultados em um gráfico, colocamos no eixo horizontal o número de grupos testados e, no eixo vertical, a soma das distâncias quadradas. Inicialmente, ao aumentar o número de grupos, a soma das distâncias diminui rapidamente, pois os dados vão sendo mais bem agrupados. No entanto, a partir de certo ponto, esse ganho se torna cada vez menor e a curva formada no gráfico começa a se estabilizar, criando um formato que lembra um cotovelo. O ponto em que essa curva faz a "dobradiça" – ou seja, onde a redução da soma das distâncias deixa de ser significativa – indica o número ideal de grupos a ser utilizado. Por exemplo, se ao passar de três para quatro grupos a diminuição da soma das distâncias é grande, mas ao passar de quatro para cinco a diferença já é pequena, o "cotovelo" da curva está em quatro grupos, sugerindo que esse é o valor mais adequado para segmentar os dados.

3.1.3. Análise de Silhueta

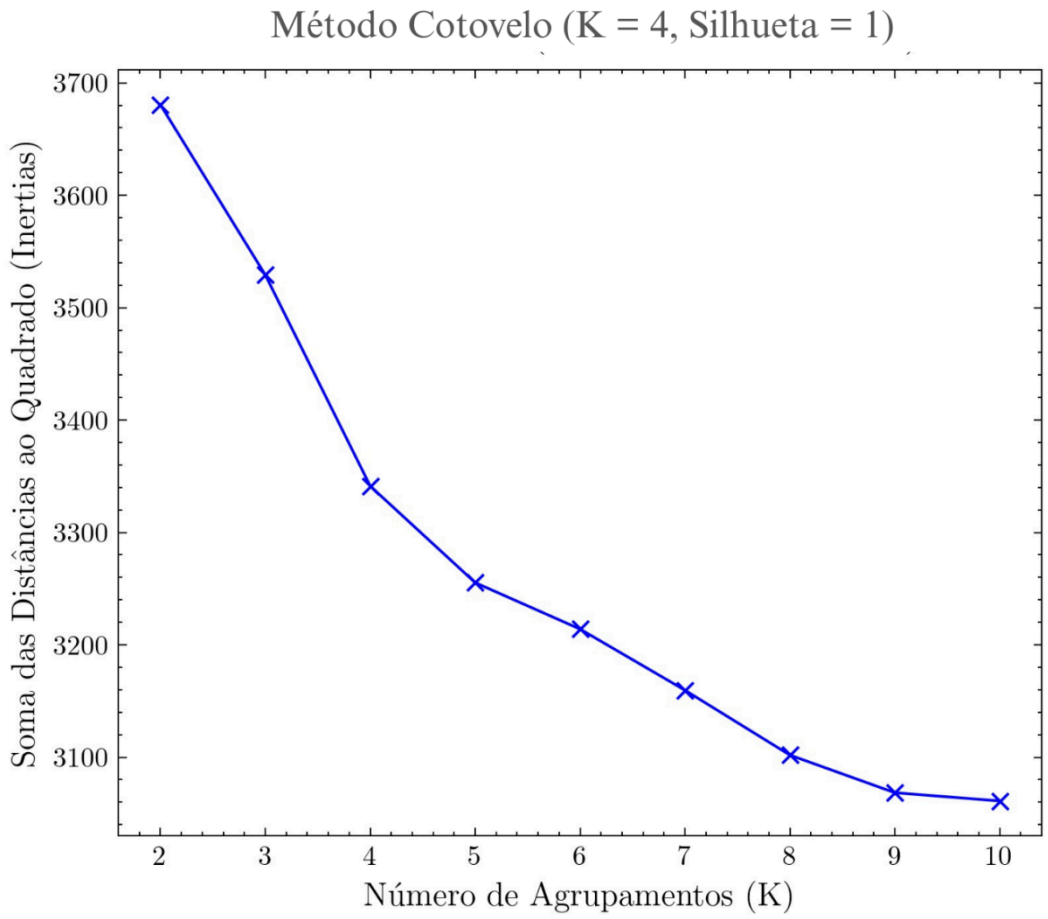
A Análise de Silhueta é uma técnica utilizada para avaliar a qualidade dos agrupamentos formados pelo algoritmo K-Means, permitindo verificar se os dados foram corretamente separados em grupos distintos. De modo didático, podemos imaginar que, após separar um conjunto de objetos em caixas, queremos saber o quão bem cada objeto realmente pertence à sua caixa e não a outra. Para isso, a Análise de Silhueta atribui a cada elemento um valor que indica o quanto ele está bem alocado em seu grupo em comparação com outros grupos. Esse valor varia de -1 a 1: quanto mais próximo de 1, melhor o ajuste do elemento ao seu grupo; valores próximos de zero indicam que o elemento está na fronteira entre dois grupos; e valores negativos sugerem que o elemento possivelmente estaria melhor em outro grupo.

Na prática, ao aplicar a Análise de Silhueta, buscamos identificar o número de grupos que apresenta a maior média desses valores para todos os elementos do conjunto de dados. Uma média

alta indica que os grupos estão bem definidos e que os elementos estão, em sua maioria, corretamente alocados. Por exemplo, se ao testar diferentes quantidades de grupos, observamos que a média da silhueta é mais alta quando usamos quatro grupos, isso sugere que essa é a melhor escolha para representar a estrutura dos dados. Assim, a Análise de Silhueta atua como uma ferramenta de validação, fornecendo uma medida objetiva e intuitiva para apoiar a decisão sobre a quantidade ideal de grupos, complementando outras técnicas como o Método do Cotovelo.

PRELO

Gráfico 1 - Número ideal de *clusters* pelo Método do Cotovelo.



Fonte: Elaborado pelos autores.

O Gráfico 1 apresenta os resultados do método cotovelo. No título do gráfico, o valor "K = 4" indica que o método identificou quatro como o número ideal de grupos para classificar os estudantes. Em termos práticos, isto significa que, ao analisar os padrões nas respostas dos 1391 estudantes ao Inventário de Consciência Metacognitiva (MAI), descobrimos que eles se distribuem naturalmente em quatro perfis distintos.

Para entender melhor este conceito, podemos usar uma analogia: imagine que estamos organizando livros em uma biblioteca. Poderíamos separá-los por inúmeros critérios e em quantidades variáveis de categorias, mas existe um número de categorias que seria mais eficiente e representativo. O Método Cotovelo nos ajuda a encontrar este número ideal, que no caso de nossa pesquisa é quatro.

No gráfico, observamos uma curva azul que forma um "cotovelo" justamente no ponto K = 4. Este formato é crucial para a interpretação: à medida que aumentamos o número de clusters (eixo horizontal), a soma das distâncias quadráticas dentro dos grupos (eixo vertical) diminui. Porém, após o ponto K = 4, essa diminuição se torna muito menos significativa, indicando que adicionar mais grupos não traria benefícios substanciais à nossa análise.

A segunda parte do título, "Silhueta = 1", refere-se a uma medida complementar que valida a qualidade dos agrupamentos formados. O coeficiente de silhueta varia de -1 a 1, sendo que:

- Valores próximos a 1 indicam que os elementos estão corretamente agrupados
- Valores próximos a 0 sugerem que os elementos estão na fronteira entre dois grupos
- Valores negativos indicam que os elementos provavelmente foram atribuídos ao grupo errado

Um valor de silhueta igual a 1 representa o máximo possível para este coeficiente, indicando um agrupamento excepcionalmente bem definido. Na prática, isto significa que os quatro perfis metacognitivos que identificamos são claramente distintos entre si, com características internas coesas. Para nosso estudo, esta descoberta é particularmente relevante, pois confirma que existem, de fato, padrões bem definidos de comportamento metacognitivo entre os estudantes de educação a distância.

A identificação e validação destes quatro perfis metacognitivos têm implicações diretas para o design de estratégias educacionais personalizadas. Ao confirmar que existem quatro maneiras predominantes pelas quais os estudantes abordam seu processo de aprendizagem na EAD, podemos desenvolver abordagens pedagógicas específicas que atendam às necessidades particulares de cada grupo.

Nas seções seguintes do artigo, detalhamos as características de cada um destes quatro perfis e discutimos como instituições de ensino podem implementar estratégias direcionadas para melhorar a experiência de aprendizagem e reduzir as taxas de evasão, considerando as particularidades metacognitivas de cada grupo identificado.

3.2. Análise Metacognitiva com *t-SNE*

A análise com t-SNE (*t-Distributed Stochastic Neighbor Embedding*) foi utilizada para visualizar como os estudantes se organizam naturalmente com base em suas respostas ao Inventário de Consciência Metacognitiva (MAI). Imagine que temos 52 perguntas complexas sobre estratégias de aprendizagem, e cada resposta é como uma coordenada em um espaço multidimensional difícil de entender. O t-SNE atua como um "tradutor visual", transformando essas 52 dimensões em um mapa bidimensional simples, onde podemos observar padrões que seriam invisíveis em tabelas ou gráficos convencionais.

Essa técnica funciona preservando as relações de proximidade entre os dados. Estudantes com respostas semelhantes no questionário aparecem próximos no gráfico, enquanto aqueles com estratégias muito diferentes ficam distantes. É como organizar um grupo de pessoas em uma sala: aquelas com interesses parecidos tendem a se agrupar, enquanto as com preferências únicas ocupam espaços isolados.

No gráfico gerado (Gráfico 2), identificamos quatro grupos principais, cada um representando um perfil metacognitivo distinto. O Perfil 1 (Roxo) é o mais numeroso e disperso, sugerindo que esses estudantes adotam múltiplas estratégias de aprendizagem, mas sem um padrão definido. Essa variedade pode indicar flexibilidade, mas também falta de método consistente. O Perfil 2 (Azul) forma um grupo mais coeso, mostrando respostas homogêneas - aqui estão os estudantes que planejam sistematicamente seus estudos, revisam regularmente o conteúdo e ajustam suas estratégias conforme necessidades específicas.

Já o Perfil 3 (Verde) apresenta uma sobreposição parcial com os dois primeiros, revelando uma transição entre comportamentos. Esses estudantes demonstram características tanto do Perfil 1 quanto do 2, mas com uma tendência maior à autorreflexão após tarefas. Por fim, o Perfil 4 (Amarelo) destaca-se por sua completa separação dos demais, indicando estratégias metacognitivas únicas - como o uso exclusivo de métodos não convencionais ou dificuldades na interpretação das próprias práticas de estudo.

A disposição espacial desses grupos no gráfico não apenas confirma a existência de perfis distintos, mas também revela relações entre eles. A proximidade entre os Perfis 1 e 2, por exemplo, sugere que alguns estudantes podem migrar entre esses grupos conforme desenvolvem habilidades metacognitivas. Já o isolamento do Perfil 4 reforça a necessidade de intervenções personalizadas para esse grupo, que provavelmente enfrenta maiores desafios na autorregulação da aprendizagem.

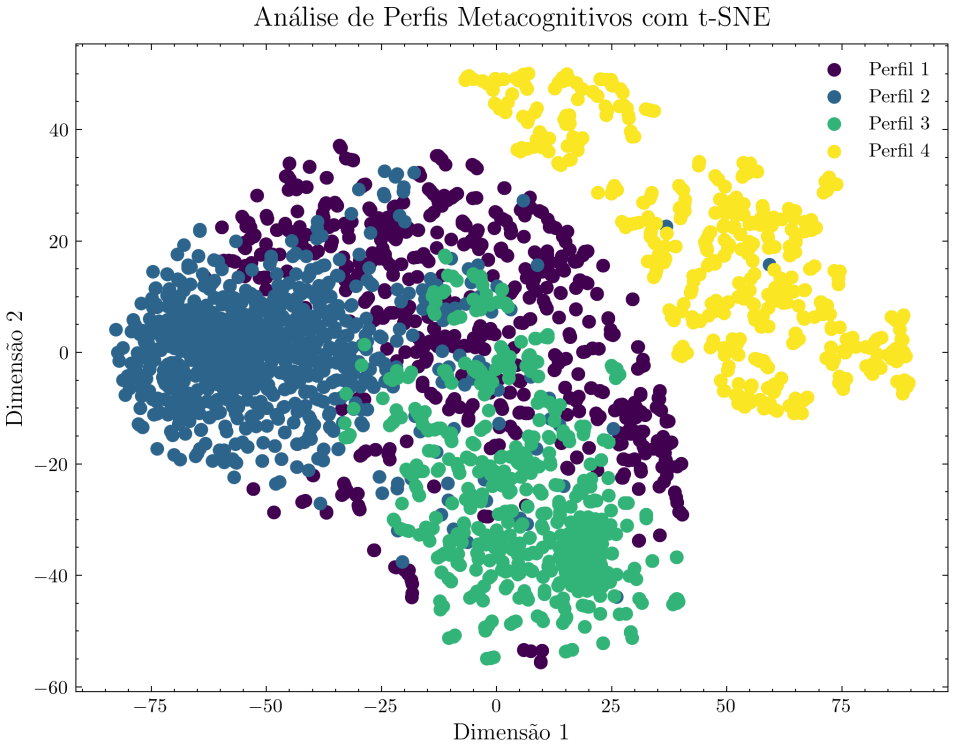
Para incrementar a compreensão do resultado obtido com o t-SNE, segue um resumo dos achados:

- **Perfil 1 (Roxo):** Este grupo possui o maior número de pontos, indicando uma prevalência deste perfil metacognitivo entre os estudantes analisados. A dispersão espacial desses pontos sugere uma diversidade interna significativa dentro do grupo, indicando variações nos comportamentos metacognitivos, mesmo dentro de um perfil comum. Essa diversidade pode refletir diferentes graus de autorregulação e estratégias de aprendizagem adotadas pelos estudantes.
- **Perfil 2 (Azul):** Embora este perfil se sobreponha parcialmente ao Perfil 1, é mais coeso e agrupado distintamente. Esta coesão pode indicar uma maior consistência nas respostas metacognitivas e um conjunto mais homogêneo de estratégias de aprendizagem entre os estudantes deste grupo. A sobreposição com o Perfil 1 sugere semelhanças nos comportamentos metacognitivos desses dois grupos, embora com distinções claras que justificam sua separação.
- **Perfil 3 (Verde):** Este agrupamento é separado do Perfil 1, com alguma sobreposição com o Perfil 2, o que pode indicar estratégias de aprendizagem compartilhadas ou características metacognitivas entre esses dois perfis. No entanto, a separação do Perfil 1 destaca diferenças fundamentais em suas abordagens metacognitivas de aprendizagem.

PRELO

- **Perfil 4 (Amarelo):** Completamente separado dos outros grupos, este perfil demonstra um conjunto distinto de características metacognitivas. A clara distinção espacial no gráfico *t-SNE* reforça a singularidade deste grupo, possivelmente refletindo estratégias de aprendizagem únicas ou uma abordagem metacognitiva distintamente diferente em comparação aos outros perfis.

Gráfico 2 - Distribuição dos perfis metacognitivos por t-SNE.



Fonte: Elaborado pelos autores.

3.3. Análise de Componentes Principais (PCA) e Perfis Analisados

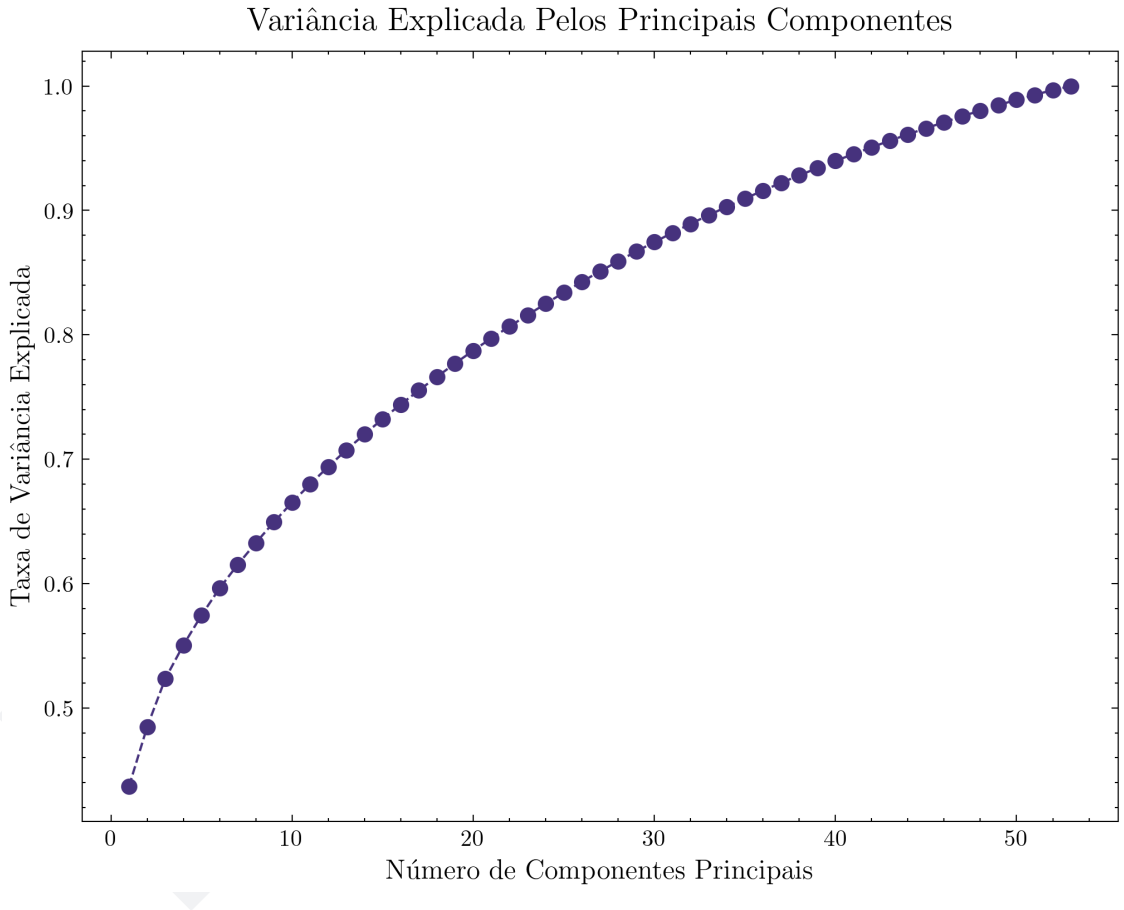
A Análise de Componentes Principais (PCA – *Principal component analysis*) pode ser compreendida como uma ferramenta que ajuda a enxergar, de forma clara e simplificada, quais características realmente diferenciam os estudantes em relação às suas estratégias metacognitivas. Imagine que cada estudante respondeu a 52 perguntas do Inventário de Consciência Metacognitiva, o que gera uma enorme quantidade de informações e torna difícil perceber, a olho nu, quais respostas de fato separam os perfis de aprendizagem. O PCA atua como uma espécie de "filtro inteligente", reduzindo essa complexidade ao identificar quais perguntas mais contribuem para distinguir os grupos formados.

No contexto deste estudo, o PCA foi aplicado após a identificação dos quatro perfis metacognitivos, justamente para revelar quais itens do questionário eram mais relevantes para diferenciar esses grupos. O método analisa a variação das respostas entre os centros (ou centróides)

de cada grupo. Quando uma pergunta apresenta grande diferença nas respostas médias entre os perfis, ela é considerada um componente principal, ou seja, um fator-chave para explicar as diferenças comportamentais entre os estudantes.

Os resultados mostraram que, dentre as 52 perguntas do MAI, dez se destacaram como as mais importantes para caracterizar os perfis metacognitivos. Essas perguntas envolvem, por exemplo, a elaboração de perguntas antes de estudar, o uso de imagens e diagramas para aprender, a definição de metas antes de iniciar uma tarefa e a autoavaliação após concluir um objetivo. Assim, o PCA permite visualizar, por meio de gráficos, como essas perguntas selecionadas capturam as principais diferenças entre os grupos, facilitando a interpretação das dinâmicas metacognitivas.

Gráfico 3 - Importância dos componentes principais (PCA) nas diferenças entre perfis.



Fonte: Elaborado pelos autores.

Para entender melhor as diferenças entre os grupos de estudantes, focamos nas perguntas do Inventário de Consciência Metacognitiva (MAI) que mais variam entre os grupos. As perguntas com maior diferença nos centros foram selecionadas, pois indicam onde os grupos mais se

distinguem. Isso nos permitiu identificar os 10 itens do MAI que melhor revelam as diferenças nos comportamentos dos estudantes.

a) Procuro elaborar perguntas sobre o material antes de começar a estudar.

b) Procuro analisar a utilidade das estratégias enquanto estudo.

c) Tenho propósitos específicos para cada estratégia que uso.

d) Sei quando cada estratégia que uso será mais eficaz.

e) Costumo usar imagens e diagramas para me ajudar a entender e aprender.

f) Quando finalizo uma tarefa, me pergunto o quão bem cumpri meus objetivos.

g) Estou ciente das estratégias de aprendizagem que uso quando estou estudando.

h) Encontro e uso estratégias de aprendizagem úteis automaticamente.

i) Costumo definir metas específicas antes de começar uma tarefa.

j) Costumo parar regularmente para verificar minha compreensão.

Subsequentemente, foram *plotados* gráficos de barras (**figura 1**) para representar visualmente os padrões comportamentais dos *clusters* criados em relação a essas perguntas-chave.

PRELO

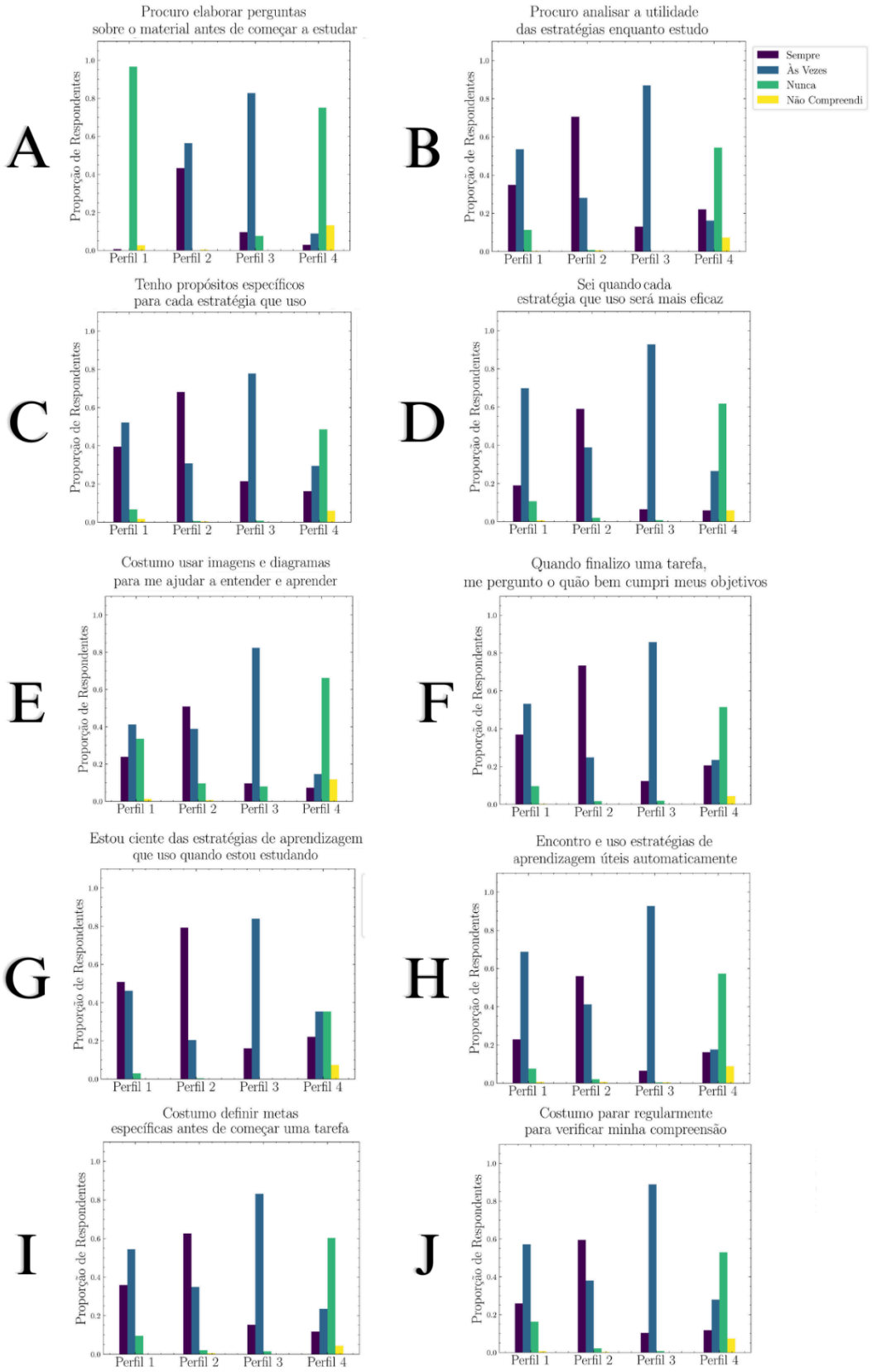


Figura 1 - Respostas dos estudantes às perguntas-chave do MAI.

Fonte: Elaborado pelos autores.

Foi então possível identificar quatro *clusters* que agruparam tipos distintos de comportamentos dos estudantes conforme avaliados pelo Inventário de Consciência Metacognitiva (MAI). Para esta análise, buscamos avaliar, dentro dos dados *plotados*, as barras com uma maioria esmagadora de participantes de um *cluster* exibindo respostas consistentes. Quando a grande maioria dos indivíduos do mesmo *cluster* fornecia respostas idênticas a uma pergunta, podíamos assumir que essa resposta definia o *cluster*, estabelecendo uma característica-chave daquele tipo particular de estudante. Para este conjunto de dados obtido a partir da coleta de dados mencionada neste estudo, caracterizamos os *clusters* da seguinte maneira:

- ***Cluster* 1:** Compreende estudantes que nunca formulam perguntas antes de estudar, às vezes sabem a eficácia de cada estratégia que empregam e às vezes encontram e utilizam estratégias de aprendizagem úteis automaticamente. Compõem 21.6% da base de respondentes.
- ***Cluster* 2:** Compreende estudantes que estão consistentemente cientes das estratégias de aprendizagem que usam ao estudar e sempre questionam o quanto atingiram suas metas ao completar tarefas. Compõem 54.7% da base de respondentes.
- ***Cluster* 3:** Compreende estudantes que respondem às vezes de forma consistente à grande maioria das perguntas. Compõem 18.8% da base de respondentes.
- ***Cluster* 4:** Compreende estudantes que nunca formulam perguntas sobre o material antes de estudar e nunca utilizam imagens e diagramas para ajudar na compreensão e aprendizagem. Além disso, podemos observar que o Cluster 3 tende a agrupar estudantes que não compreenderam as perguntas feitas. Compõem 4.9% da base de respondentes.

4. Discussão

Os resultados obtidos através da aplicação do algoritmo *K-Means*, juntamente com as análises de *Elbow* e Silhueta, forneceram *insights* significativos sobre a estrutura subjacente dos dados metacognitivos coletados. A interpretação dos resultados revela que a estratégia de quatro *clusters* é a mais adequada para descrever a variação nos comportamentos dos estudantes em relação às perguntas do Inventário de Consciência Metacognitiva (MAI).

Este agrupamento revelou quatro perfis metacognitivos distintos entre os participantes do estudo. Vamos analisar esses *clusters* e suas principais características:

- ***Cluster* 0:** Compreende estudantes que raramente formulam perguntas antes de estudar, ocasionalmente estão cientes da eficácia das estratégias que utilizam e às

vezes encontram e utilizam automaticamente estratégias de aprendizagem úteis. Isso sugere um nível variável de autorregulação e autoconsciência em relação às estratégias de aprendizagem.

- ***Cluster* 1:** Estudantes neste cluster demonstram um alto grau de autorregulação e autoconsciência. Eles estão consistentemente cientes das estratégias de aprendizagem que empregam durante o estudo e constantemente autoavaliam seu alcance de metas ao concluir tarefas. Este perfil metacognitivo indica estudantes que abordam a aprendizagem de maneira deliberada e consciente.

- ***Cluster* 2**: Compreende estudantes que variam em suas respostas à maioria das perguntas. Esta variabilidade sugere que esses estudantes não seguem um padrão metacognitivo consistente e podem adotar diferentes estratégias de aprendizagem em momentos distintos.

- ***Cluster* 3:** Notavelmente, este cluster inclui estudantes que nunca formulam perguntas sobre o material antes de estudar e nunca utilizam imagens e diagramas para auxiliar na compreensão e aprendizagem. Além disso, esses estudantes demonstraram uma compreensão limitada das perguntas feitas no questionário. Este cluster pode representar um grupo que poderia se beneficiar de intervenções específicas para desenvolver suas habilidades metacognitivas e estratégias de aprendizagem.

A análise adicional através da medida de variância dos centróides revelou que as dez perguntas selecionadas são cruciais na definição dos perfis metacognitivos. Estas perguntas cobrem tópicos como formulação de perguntas antes de estudar, análise da utilidade das estratégias, estabelecimento de metas específicas e avaliação da eficácia das estratégias. A seleção dessas perguntas permitiu uma compreensão mais profunda da variação nos comportamentos dos estudantes em relação a aspectos cruciais da autorregulação e metacognição.

O uso do *t-SNE* para analisar os perfis metacognitivos ofereceu uma compreensão visual profunda das relações entre os estudantes, destacando a complexidade dos comportamentos de aprendizagem e as nuances entre os perfis identificados. Esta visualização enfatiza a heterogeneidade dos estudantes em cursos de Educação a Distância (EAD) e sublinha a importância de abordagens educacionais personalizadas capazes de atender às necessidades específicas de diferentes grupos de aprendizes. Através desta análise, torna-se evidente a necessidade de estratégias de ensino adaptativas que considerem a diversidade metacognitiva dos estudantes para otimizar o processo de ensino-aprendizagem na EAD.

No geral, esta análise de *clusters* fornece *insights* ricos sobre os comportamentos metacognitivos dos estudantes de ensino a distância. Os resultados destacam a heterogeneidade

nas abordagens de aprendizagem e indicam áreas onde intervenções podem ser valiosas. Compreender essas diferenças nos perfis metacognitivos dos estudantes pode permitir estratégias de ensino mais direcionadas, suporte personalizado e o desenvolvimento de recursos educacionais que atendam às necessidades específicas de cada grupo. Identificar estudantes em risco, possivelmente dentro do *Cluster* 3, também abre caminho para a implementação de medidas de suporte preventivo, visando melhorar a experiência de aprendizagem e reduzir as taxas de abandono.

No entanto, é essencial lembrar que os resultados da análise de *clusters* representam um instantâneo em um determinado momento. À medida que os estudantes progridem em seus cursos, seus comportamentos metacognitivos podem evoluir. Portanto, as instituições de ensino a distância devem considerar a aplicação contínua de técnicas de análise de dados para monitorar e responder às mudanças nos perfis metacognitivos dos estudantes, contribuindo para a melhoria contínua da qualidade da educação a distância.

4.1. Evasão e os Perfis Metacognitivos

A evasão em cursos de Educação a Distância (EAD) permanece como um desafio recorrente, afetando diretamente a eficácia do processo de ensino e o êxito dos estudantes. Nesta pesquisa, investigou-se a relação entre os perfis metacognitivos dos alunos e suas taxas de evasão, com o objetivo de compreender como diferentes níveis de autopercepção e estratégias de aprendizagem influenciam a permanência nos cursos. A análise metacognitiva identificou perfis distintos entre os estudantes, caracterizados por diferentes graus de consciência sobre seus próprios processos de pensamento e pela capacidade de autorregulação. Em outras palavras, alguns alunos demonstraram maior habilidade para planejar, monitorar e avaliar seu próprio aprendizado.

Estudos anteriores já apontam uma correlação positiva entre essas habilidades metacognitivas e o sucesso acadêmico, indicando que estudantes mais conscientes de seu processo de aprendizagem tendem a apresentar melhores resultados e menor propensão ao abandono. Ademais, pesquisas recentes indicam que tecnologias baseadas em inteligência artificial, como os sistemas adaptativos de recomendação e tutoria, têm potencial para personalizar ainda mais a experiência educacional online, promovendo o engajamento contínuo e reduzindo taxas de evasão (Zawacki-Richter *et al.*, 2019).

Estudos anteriores também estabeleceram uma correlação positiva entre metacognição e sucesso acadêmico, sugerindo que uma maior habilidade metacognitiva está associada a melhores resultados de aprendizagem e a uma menor propensão ao abandono (Flavell, 1979; Karaoglan Yilmaz; Yilmaz, 2020; Zimmerman, 2002). Estudantes com alta consciência metacognitiva tendem a empregar estratégias de aprendizagem mais eficazes, ajustando-as conforme necessário para otimizar a compreensão e retenção de conteúdo. Em contraste,

estudantes com menor consciência metacognitiva podem não reconhecer quando suas estratégias de aprendizagem são ineficazes, levando à frustração e, potencialmente, à desmotivação e abandono. (Karaoglan Yilmaz; Yilmaz, 2020)

Os perfis metacognitivos identificados neste estudo sugerem uma relação direta entre a capacidade dos estudantes de refletir sobre sua própria aprendizagem e sua persistência em cursos EAD. Especificamente, perfis caracterizados por menor consciência metacognitiva e estratégias de aprendizagem menos eficientes mostraram-se mais vulneráveis à desmotivação. Esta desmotivação, frequentemente exacerbada pela falta de *feedback* imediato e suporte personalizado típico do ambiente EAD, pode precipitar a decisão de abandonar o curso.

Diante desses achados, torna-se imperativo adotar abordagens pedagógicas que não apenas promovam a consciência metacognitiva, mas também forneçam suporte adaptativo aos estudantes com base em seus perfis de aprendizagem. Estratégias como feedback personalizado, orientação específica para o desenvolvimento de habilidades de estudo e a criação de comunidades de aprendizagem podem ser particularmente eficazes em engajar estudantes com diferentes perfis metacognitivos, reduzindo assim as taxas de abandono.

Portanto, reconhecer e responder às necessidades específicas dos estudantes com base em sua consciência e habilidades metacognitivas é crucial para o design instrucional na EAD. Implementar intervenções focadas na metacognição pode não apenas melhorar o sucesso acadêmico, mas também aumentar a retenção dos estudantes, abordando um dos maiores desafios enfrentados pela educação online contemporânea. Este estudo reforça a necessidade de esforços contínuos para investigar as dinâmicas metacognitivas em ambientes EAD, visando aprimorar estratégias que promovam uma experiência de aprendizagem mais envolvente e reduzam a propensão ao abandono.

4.2. Estratégias Neuroeducacionais baseadas nos Perfis Metacognitivos

A análise dos perfis metacognitivos dos estudantes em cursos de Educação a Distância (EAD) revela uma oportunidade única de desenvolver estratégias pedagógicas personalizadas, voltadas não apenas para otimizar o processo de ensino-aprendizagem, mas também para reduzir as taxas de evasão. Considerando os quatro perfis metacognitivos identificados, é imperativo que as técnicas de design instrucional e o conhecimento neurocientífico sejam estrategicamente empregados para atender às necessidades específicas de cada grupo.

- **Estratégias para o Perfil 1:** Para estudantes que mostram maior dispersão em suas estratégias de aprendizagem, sugere-se a implementação de recursos educacionais que favoreçam a autoexploração e o autoconhecimento. Técnicas como mapas mentais e diários de aprendizagem podem ser úteis para estimular a

reflexão e a consciência sobre seus próprios processos de pensamento e estratégias de estudo.

- **Estratégias para o Perfil 2:** Estudantes com um perfil de aprendizagem mais coeso e específico podem se beneficiar de técnicas de design instrucional que promovam a aplicação prática do conhecimento, como estudos de caso e simulações. Essas atividades, baseadas em evidências neurocientíficas sobre a aplicação e consolidação do conhecimento, podem encorajar o engajamento e aprofundar a compreensão do conteúdo.

- **Estratégias para o Perfil 3:** Para estudantes que demonstram dificuldades em estratégias de aprendizagem autodirigidas, é crucial reforçar a orientação sobre técnicas de estudo e planejamento. Recursos como tutoriais interativos sobre métodos de estudo eficazes, sessões de *coaching* e workshops sobre gestão do tempo e estabelecimento de objetivos de aprendizagem podem ser extremamente valiosos.

- **Estratégias para o Perfil 4:** O perfil caracterizado por estratégias metacognitivas únicas e talvez não convencionais requer uma abordagem altamente personalizada. Ferramentas de aprendizagem adaptativas, que ajustam o conteúdo e o nível de desafio com base no desempenho e nas respostas do estudante, podem ser particularmente eficazes. A integração de feedback personalizado e imediato também é essencial para manter esses estudantes engajados e em uma trajetória de aprendizagem positiva.

Em todos os casos, promover a autonomia do estudante é fundamental. O conhecimento neurocientífico aplicado ao design instrucional, como a importância da atenção, memória e motivação, deve guiar a criação de ambientes de aprendizagem que estimulem a curiosidade, ofereçam suporte adaptativo e reconheçam as conquistas individuais. Tais ambientes não apenas aprimoram a aprendizagem, mas também contribuem para diminuir a sensação de isolamento e a probabilidade de evasão.

5. Conclusão

Diante dos *insights* abrangentes obtidos da interseção entre Neuroeducação, análise metacognitiva e Ciência de Dados na Educação a Distância (EAD), este estudo ilumina o profundo impacto da compreensão e atendimento aos diversos perfis metacognitivos dos estudantes na melhoria de suas experiências de aprendizagem e mitigação das taxas de abandono. Através da aplicação meticulosa de técnicas de mineração de dados, incluindo agrupamento e Análise de Componentes Principais (PCA), identificamos perfis metacognitivos distintos entre os estudantes de pós-graduação. Esses perfis não apenas sublinham a variabilidade nas estratégias de aprendizagem e consciência cognitiva dos estudantes, mas também destacam o potencial de

intervenções educacionais personalizadas para melhorar significativamente o engajamento dos estudantes e o sucesso acadêmico em ambientes EAD.

A implantação de estratégias neuroeducacionais informadas pelas necessidades e características únicas de cada perfil metacognitivo sugere um caminho promissor para a personalização da aprendizagem. Ao alinhar o design instrucional com a compreensão neurocientífica e os *insights* metacognitivos, educadores e designers instrucionais podem fomentar ambientes de aprendizagem mais eficazes e envolventes. Esta abordagem não apenas aborda os desafios imediatos de retenção de estudantes e abandono, mas também avança o objetivo mais amplo de cultivar aprendizes ao longo da vida, equipados com a autoconsciência e adaptabilidade para navegar em um cenário educacional em constante evolução.

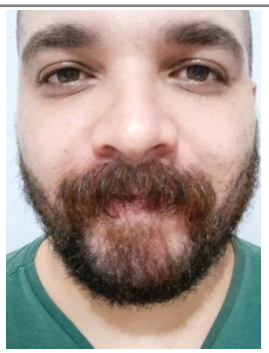
Além disso, esta pesquisa enfatiza a necessidade de inovação contínua e adaptabilidade nas estratégias educacionais para acomodar as complexidades dos processos de aprendizagem dos estudantes. A integração da Ciência de Dados na prática educacional oferece uma ferramenta promissora para descobrir e responder a essas complexidades, permitindo uma abordagem informada por dados para o design instrucional e suporte aos estudantes.

Em conclusão, este estudo contribui para o crescente corpo de conhecimento na interseção entre Neuroeducação, metacognição e Ciência de Dados, destacando o potencial transformador dessas disciplinas na melhoria da educação a distância. À medida que continuamos a explorar e entender as dinâmicas intrincadas da aprendizagem, os achados desta pesquisa sublinham a importância de uma abordagem educacional centrada no estudante e baseada em evidências. Ao abraçar a diversidade dos perfis de aprendizes e aproveitar os *insights* fornecidos pela Ciência de Dados, podemos nos aproximar da realização do pleno potencial da educação a distância como uma modalidade flexível, inclusiva e eficaz para aprendizes em todo o mundo.

Biodados e contatos dos autores

DA SILVA, V. G. R. é doutora em Neurociências e Biologia Celular pela UFPA e designer instrucional, que atua criando cursos, estratégias metodológicas, gerenciando equipes e processos de criação de materiais didáticos que otimizam a experiência de ensino-aprendizagem através de técnicas de Neurolearning e UX/LX Design na modalidade presencial e EAD. Suas áreas de interesse de estudo e pesquisa incluem neuropsicologia, psicoeducação e ciências de dados aplicadas à educação e saúde.

ORCID: https://orcid.org/0000-0002-7979-9622
Contato: +55 91 99160 4849
E-mail: vr.neuropsi@gmail.com

	SILVA, A. G. é doutorando em Fisiopatologia Médica – Neurociências pela UNICAMP, mestre em Neurociência e Comportamento (UFPA) e cientista de dados premiado. Atua na interface entre neurociências e inteligência artificial, com experiência em processamento de sinais digitais, imagens médicas, machine learning e desenvolvimento de software. Suas áreas de interesse incluem diagnóstico assistido por IA, redes neurais biologicamente plausíveis e neuromorphic hardware. ORCID: https://orcid.org/0009-0008-8384-6381 E-mail: a257073@dac.unicamp.br
	BASTOS, C. O. é mestre em Neurociências em Neurociências e Biologia Celular pela Universidade Federal do Pará. Seus interesses de pesquisa incluem aprendizado de máquina, deep learning e LLM com destaque para as a otimização de algoritmos de RAG. ORCID: https://orcid.org/0000-0003-4687-788X E-mail: bastoscaio96@gmail.com
	GOMES, B. D. é professor associado do Instituto de Ciências Biológicas da UFPA, doutor pela UFPA e pós-doutor pelo MIT. Atua em neurofisiologia, cognição, neurociência computacional e inteligência artificial, com ênfase em teoria de sistemas complexos e IA aplicada à saúde e neurociência. É pesquisador afiliado ao MIT, chefe de pesquisa em startups de IA na saúde e vice-diretor da Faculdade de Biotecnologia da UFPA. ORCID: https://orcid.org/0000-0002-3646-7847 E-mail: brunodgomes@ufpa.br

Agradecimentos

Os autores gostariam de expressar sua profunda gratidão aos líderes da equipe Estratego por sua abertura e fornecimento de dados essenciais para este estudo. Agradecimentos também aos estudantes de pós-graduação cuja participação foi crucial para o desenvolvimento desta pesquisa.

Referências Bibliográficas

ABBAD, G. *et al*. **Evasão em curso via internet: explorando variáveis explicativas.** RAE eletrônica, v. 5, n. 2, 2006.

AHMED, M. *et al*. **The k-means Algorithm: A Comprehensive Survey and Performance Evaluation.** Electronics, v. 9, n. 8, p. 1295, 2020.

ALLIPRANDINI, P. S, A. *et al.* **Estratégias de aprendizagem utilizadas por estudantes na educação a distância: implicações educacionais.** Psicologia da Educação, São Paulo , n. 38, p. 05-16, 2014.

ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO A DISTÂNCIA. **Censo EAD.BR: relatório analítico da aprendizagem a distância no Brasil 2020**. Curitiba, PR: InterSaberes, 2022.

BAHAR, A. S.; SHAPIRO, M. L. **Remembering to Learn: Independent Place and Journey Coding Mechanisms Contribute to Memory Transfer**. Journal of Neuroscience, v. 32, n. 6, p. 2191–2203, 2012.

COSENZA, R. M; GUERRA, L. **Neurociência e educação: como o cérebro aprende.** Porto Alegre: Artmed, 2011.

COSTA, R D. **Classificação dos estilos de aprendizagem baseado em sistemas inteligentes: um estudo de caso na educação mediada por tecnologia**. 81f. Tese (Doutorado em Engenharia Elétrica e de Computação) - Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal, 2020.

FERNANDEZ, A. *et al*. **SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary.** Journal of Artificial Intelligence Research, v. 61, p. 863–905, 2018.

FILATRO, A.; CAIRO, S. **Produção de Conteúdos Educacionais.** São Paulo: Ed. Saraiva Uni, 2015.

FILATRO, A. **Data Science na Educação: Presencial, a Distância e Corporativa**. São Paulo, SP: Saraiva Educação, 2021.

FILATRO, A. **Design instrucional contextualizado: educação e tecnologia**. São Paulo: Ed. SENAC, 2004.

FILATRO, A. **Design Instrucional na prática**. 1ª Ed. Pearson, 2008.

FLAVELL, J. H. **Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry**. American Psychologist, v. 34, n. 10, p. 906-911, 1979.

FRAZIER, P. I. **A Tutorial on Bayesian Optimization**. arXiv (Cornell University), 2018.

GHADDAR, B.; NAOUM-SAWAYA, J. **High dimensional data classification and feature selection using support vector machines**. European Journal of Operational Research, v. 265, n. 3, p. 993–1004, 2018.

GUTIERREZ-PACHAS, D. A. *et al*. **Supporting Decision-Making Process on Higher Education Dropout by Analyzing Academic, Socioeconomic, and Equity Factors through Machine Learning and Survival Analysis Methods in the Latin American Context.** Education Sciences, v. 13, n. 2, 154, 2023. Disponível em: https://www.mdpi.com/2227-7102/13/2/154 Acesso em: 23 mar. 2025.

JESUS, Â. M. *et al*. **Aplicando ciência de dados educacionais para avaliar a influência da programação no progresso das notas do Ensino Médio**. Revista Educação Pública, v. 21, n. 33, 2021.

KANDEL E. R. *et al*. M. **Princípios de Neurociências**. 5a Ed. Porto Alegre: Artmed, 2014.

KARAOGLAN YILMAZ, F. G.; YILMAZ, R. **Learning analytics as a metacognitive tool to influence learner transactional distance and motivation in online learning environments.** Innovations in Education and Teaching International, v. 58, n. 5, p. 575–585, 2020.

KOLB, D. A. **Learning Style Inventory: Self Scoring Test and Interpretation Booklet**. Boston, MA: McBer, 1985.

LIMA FILHO, R. N.; BRUNI, A. L. **Metacognitive Awareness Inventory: tradução e validação a partir de uma análise fatorial confirmatória**. Psicologia: Ciência e Profissão, 35(4), 1275-1293, 2015.

LINDERMAN, G. C.; STEINERBERGER, S. **Clustering with t-SNE, provably. SIAM journal on mathematics of data science**, v. 1, n. 2, p. 313–332, 2019.

MORY, E. H. **Feedback Research Review. In: JONASSEM, D. (Comp.). Handbook of Research on Educational Communications and Technology**. Mahwah: Lawrence Erlbaum. p. 745-783, 2004.

ORSUCCI, F. F; SALA, N. **Neuroscience and Technology Transform the Educational Ecosystem.** New York: Nova Medical and Health, 2022.

PELIKAN, E. R. *et al.* **Distance learning in higher education during COVID-19: The role of basic psychological needs and intrinsic motivation for persistence and procrastination–a multi-country study.** PLOS ONE, v. 16, n. 10, p. e0257346, 2021.

PINTO, L. F. G. **Teorias De Aprendizagem Aplicadas ao E-learning: Uma Abordagem da Teoria Cognitiva De Aprendizagem Multimídia**. Anais do Congresso Internacional de Educação e Tecnologias, 2020.

RABELO, D. S. S. *et al.* **Utilização de técnicas de mineração de dados educacionais para predição de desempenho de alunos EaD em ambientes virtuais de aprendizagem.** Anais do Simpósio Brasileiro de Informática na Educação (SBIE), Recife-PE, 2017.

RAJU, V. N. G. *et al*. **Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification**. Disponível em: https://ieeexplore.ieee.org/document/9214160 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 729-735. IEEE, 2020.

RAMOS, A. S. F. **Dados recentes da neurociência fundamentam o método "Brain-based learning"**. Rev. psicopedag., São Paulo , v. 31, n. 96, p. 263-274, 2014.

RIVAS, S. F. *et al.* **Metacognitive strategies and development of critical thinking in higher education**. Frontiers in Psychology, v. 13, n. 1, 2022.

ROMERO, C.; VENTURA, S. **Data mining in education**. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, v. 3, n. 1, p. 12-27, 2013.

SANTOS, L. R. dos *et al* . O Ensino Remoto Emergencial na Perspectiva da Metacognição: Análise da Percepção de Alunos de um Curso Técnico em Enfermagem. EaD em Foco, [S. l.], v. 11, n. 2, 2021. Disponível em: https://eademfoco.cecierj.edu.br/index.php/Revista/article/view/1260 Acesso em: 17 mai. 2024.

SANTOS, L. R. DOS; PEIXOTO, M. A. P. **Análise do inventário de consciência metacognitiva de alunos do curso técnico em enfermagem.** Research, Society and Development, v. 10, n. 12, p. e62101220019, 2021.

SAPUTRA, D. M. *et al*. **Effect of Distance Metrics in Determining K-Value in K-Means Clustering Using Elbow and Silhouette Method**. Proceedings of the Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019), 2020.

SCHRAW, G.; DENNISON, R. S. **Assessing metacognitive awareness.** Contemporary Educational Psychology, 19, (4), 460–475, 1994.

SILVA JÚNIOR, S. L. *et al*. **The interface of neuroscience, education and technology: enhancing learning in the twenty-first century.** Revista Aracê, São José dos Pinhais, v. 6, n. 2, p. 1419–1430, 2024. Disponível em: https://periodicos.newsciencepubl.com/arace/article/view/732 Acesso em: 23 mar. 2025.

TOMMI RAIJ *et al.* **Parallel input makes the brain run faster.** NeuroImage, v. 40, n. 4, p. 1792–1797, 2008.

WOLD, S. *et al*. **Principal component analysis.** Chemometrics and Intelligent Laboratory Systems, v. 2, n. 1-3, p. 37–52, 1987.

ZAWACKI-RICHTER, O. *et al*. **Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where are the Educators?** International Journal of Educational Technology in Higher Education, v. 16, n. 39, p. 1–27, 2019. Disponível em: https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-019-0171-0 Acesso em: 23 mar. 2025.

ZHANG, X. *et al*. **A Review of Data Mining in Personalized Education: Current Trends and Future Prospects.** arXiv preprint arXiv:2402.17236, 2024. Disponível em: https://arxiv.org/abs/2402.17236 Acesso em: 23 mar. 2025.

COMO CITAR ESTE TRABALHO

ABNT: DA SILVA, V. G. R. *et al.* Análise de Perfis Metacognitivos para Redução da Evasão em EAD: uma Abordagem Baseada em Ciência de Dados. **EaD em Foco**, v. 15, n. 1, e2377, 2025. doi: https://doi.org/10.18264/eadf.v15i1.2377