



Revista
Educar Mais

Mapeamento dos perfis comportamentais de alunos em um MOOC brasileiro

Mapping the behavioral profiles of students in a brazilian MOOC

Mapeo de los perfiles de comportamiento de los estudiantes en un MOOC brasileño

Vanessa Faria de Souza¹ 

RESUMO

Embora o auge dos cursos MOOC (Massive Open Online Courses) tenha ocorrido em 2012, muitas Instituições de Ensino ainda apostam nesse formato educacional como estratégia para a disseminação de conhecimento especializado a custos acessíveis, especialmente em um contexto no qual o ensino remoto se tornou cada vez mais relevante. Nesse sentido, pesquisas indicam que, em 2018, mais de 11 mil MOOC foram ofertados por cerca de 900 universidades, evidenciando a elevada demanda por esse modelo educacional. Todavia, diversas questões ainda permanecem em aberto no que se refere aos MOOC, destacando-se o baixo envolvimento dos participantes e as elevadas taxas de desistência, que podem alcançar aproximadamente 90%. Esses fatores têm fomentado pesquisas voltadas à personalização dos ambientes e conteúdos dos MOOC. Nesse contexto, o mapeamento dos perfis dos alunos que realizam esses cursos configura-se como uma tarefa relevante. Assim, este estudo tem como objetivo mapear os perfis de alunos de um MOOC da área de Química, ofertado por uma plataforma brasileira, com foco nos objetivos desses estudantes ao cursarem esse tipo de formação. Para a realização do mapeamento, foi utilizado o algoritmo de agrupamento K-means, a partir do qual foram identificados quatro perfis comportamentais predominantes: Engajados, Estratégicos, Inativos e Instrumentais. Os resultados obtidos neste estudo podem subsidiar o desenvolvimento de estratégias voltadas à melhoria dos MOOC ofertados pela plataforma, a partir do conhecimento dos perfis dos estudantes, contemplando ações tanto para a inibição de comportamentos inadequados quanto para o incentivo a comportamentos desejáveis.

Palavras-chave: MOOC; Mapeamento de Perfil; Perfil Comportamental; K-means.

ABSTRACT

Although the peak of Massive Open Online Courses (MOOCs) occurred in 2012, many educational institutions continue to invest in this educational format as a strategy for disseminating specialized knowledge at accessible costs, especially in a context in which remote education has become increasingly relevant. In this regard, studies indicate that in 2018 more than 11,000 MOOCs were offered by approximately 900 universities, highlighting the strong demand for this educational model. However, several issues regarding MOOCs remain open, particularly low learner engagement and high dropout rates, which can reach approximately 90%. These factors have encouraged research aimed at personalizing MOOC environments and content. In this context, mapping the profiles of students enrolled in these courses represents a significant task. Thus, this study aims to map the profiles of students enrolled in a Chemistry MOOC offered by a Brazilian platform, focusing on learners' objectives when participating in this type of course. The K-means clustering algorithm was used to perform the profiling, resulting in the identification of four predominant behavioral profiles: Engaged, Strategic, Inactive, and Instrumental. The results of this study may support the development of strategies to improve the MOOCs

¹ Bacharel em Sistemas de Informação, Licenciada em Computação, Mestra em Informática, Doutora em Informática na Educação e Docente do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS), Ibirubá/RS – Brasil. E-mail: vanessa.souza@vacaria.ifrs.edu.br

offered by the platform, based on knowledge of student profiles, including actions to inhibit undesirable behaviors and encourage desirable ones.

Keywords: MOOCs; Profile Mapping; Behavioral Profile; K-means.

RESUMEN

Aunque el auge de los cursos MOOC (Massive Open Online Courses) ocurrió en 2012, muchas instituciones educativas continúan apostando por este formato como una estrategia para la difusión de conocimiento especializado a costos accesibles, especialmente en un contexto en el que la educación a distancia se ha vuelto cada vez más relevante. En este sentido, investigaciones indican que en 2018 se ofrecieron más de 11.000 MOOCs en aproximadamente 900 universidades, lo que evidencia la alta demanda de este modelo educativo. No obstante, aún persisten diversas cuestiones abiertas relacionadas con los MOOCs, entre las que destacan el bajo nivel de compromiso de los estudiantes y las elevadas tasas de abandono, que pueden alcanzar alrededor del 90%. Estos factores han impulsado investigaciones orientadas a la personalización de los entornos y contenidos de los MOOCs. En este contexto, el mapeo de los perfiles de los estudiantes que participan en estos cursos se configura como una tarea relevante. De este modo, el presente estudio tiene como objetivo mapear los perfiles de los estudiantes de un MOOC del área de Química, ofrecido por una plataforma brasileña, con énfasis en los objetivos de los alumnos al realizar este tipo de curso. Para llevar a cabo el mapeo, se utilizó el algoritmo de agrupamiento K-means, mediante el cual se identificaron cuatro perfiles conductuales predominantes: Comprometidos, Estratégicos, Inactivos e Instrumentales. Los resultados obtenidos en este estudio pueden servir de base para el desarrollo de estrategias orientadas a la mejora de los MOOCs ofrecidos por la plataforma, a partir del conocimiento de los perfiles de los estudiantes, contemplando acciones tanto para inhibir comportamientos inadecuados como para incentivar comportamientos deseables.

Palabras clave: MOOCs; Mapeo de perfiles; Perfil de Comportamiento; K-means.

1. INTRODUÇÃO

Os Massive Open Online Courses (MOOC) têm recebido grande atenção desde seu surgimento, em meados de 2008, sobretudo em função de características como flexibilidade, amplo acesso a materiais e aulas elaborados por especialistas, além da facilidade de ingresso. De acordo com Shah (2021), o número e a diversidade de MOOCs continuam em expansão desde sua consolidação como modalidade educacional, ultrapassando 220 milhões de matrículas acumuladas em plataformas globais e envolvendo milhares de cursos ofertados por universidades e empresas ao redor do mundo. Apesar do crescimento expressivo e da ampliação do acesso à educação, persistem desafios significativos associados a esse modelo. Entre as problemáticas mais recorrentes na literatura destacam-se as elevadas taxas de evasão, que frequentemente permanecem acima de 80%, especialmente em cursos autoinstrucionais e de curta duração, evidenciando lacunas relacionadas ao envolvimento, à motivação e ao suporte ao estudante.

Na tentativa de mitigar os problemas vinculados a esse modelo educacional, diversas pesquisas têm sido desenvolvidas com o objetivo de identificar os principais perfis comportamentais dos alunos que realizam MOOC. Como exemplos, podem-se citar os estudos conduzidos por Rodrigues *et al.* (2016), Tan *et al.* (2018), Mareca e Bordel (2019) e Shi, Peng e Wang (2017). Esses estudos têm se intensificado, uma vez que, em uma sociedade cada vez mais informatizada, a oferta de serviços customizados, personalizados e adaptáveis — embora represente um desafio — mostra-se vantajosa, pois contribui para a melhoria de serviços e funcionalidades em diferentes sistemas. Nesse contexto, o conhecimento do perfil do usuário constitui uma etapa fundamental, e a modelagem de perfis configura-se como um campo relevante, cujo objetivo é fornecer uma representação abstrata de aspectos relacionados às características do usuário.

No campo educacional, a modelagem do perfil do aluno pode subsidiar a tomada de decisões em diferentes níveis, como, por exemplo, na gestão de problemas relacionados ao baixo desempenho e à evasão; na oferta de orientações e recomendações mais adequadas aos estudantes; e na definição de recursos de aprendizagem mais adaptáveis, de acordo com o perfil do discente. Nesse sentido, em cursos de *e-learning*, particularmente os MOOC, essa abordagem pode contribuir para a personalização dos ambientes de aprendizagem e das plataformas que ofertam esses cursos, resultando em um maior número de alunos concluintes, melhores experiências de aprendizado e maior conhecimento agregado. Ademais, possibilita uma compreensão mais aprofundada sobre a forma como os estudantes se comportam na condução de seu próprio aprendizado, o que pode apoiar a elaboração de conteúdos mais adequados a diferentes públicos-alvo.

O perfil do aluno pode ser entendido como uma estrutura que reúne informações diretas e indiretas sobre seu comportamento, abrangendo interesses, preferências, dados pessoais e habilidades (Ferreira-Satler *et al.*, 2012). Frequentemente, o conceito de perfil está diretamente relacionado ao contexto analisado. No âmbito dos MOOC, por exemplo, o foco recai sobre a interação do aluno com a plataforma, considerando aspectos como tempo e duração de conexão, navegação no ambiente virtual. Diante desse cenário, este estudo tem como objetivo mapear os perfis comportamentais dos alunos de um MOOC da área de Química, oferecido por uma plataforma brasileira. A definição desses perfis é relevante para a obtenção de uma compreensão mais ampla acerca do comportamento dos estudantes ao longo da realização de um MOOC, podendo contribuir para a personalização do ambiente educacional, o desenvolvimento de cursos mais atrativos, o fortalecimento do envolvimento discente e a melhoria dos índices de conclusão. Além disso, tal compreensão pode dar suporte à implementação de mecanismos e diretrizes voltados à inibição de práticas inadequadas durante a realização de MOOC, como plágio e/ou cópia de respostas em atividades avaliativas.

A definição dos perfis de alunos em uma plataforma MOOC não constitui uma tarefa trivial. Seu desenvolvimento exigiu a realização de uma pesquisa abrangente, iniciando pela conceituação de padrões comportamentais de aprendizagem, incluindo as principais técnicas para o mapeamento de perfis de alunos, até a análise de exemplos que demonstram como essas técnicas podem ser aplicadas nesse contexto. Essa investigação foi sistematizada nas Seções 2 e 3 deste estudo. Posteriormente, na Seção 4, é apresentada a abordagem de mapeamento de perfis adotada, destacando-se as técnicas empregadas no mapeamento dos perfis dos alunos do MOOC analisado. Em seguida, os resultados obtidos são apresentados na Seção 5, e, por fim, as considerações finais são discutidas na Seção 6.

No que se refere à conceituação de perfis de alunos, é relevante destacar que, na pesquisa realizada, não foi identificada uma publicação que estabeleça uma diferenciação clara entre termos como perfil comportamental, perfil de aprendizagem ou estilos de aprendizagem no contexto educacional. Observou-se que a terminologia adotada tende a depender do aspecto que se deseja analisar. Há, por exemplo, estudos que investigam perfis de engajamento, analisando o nível de comprometimento do aluno com o curso, como em Rodrigues *et al.* (2016), enquanto outros trabalhos buscam mapear perfis relacionados à busca por ajuda, como em Finney *et al.* (2018). À vista disso, compreende-se que a terminologia utilizada está condicionada ao contexto da investigação. Nesse sentido, a análise dos perfis dos alunos desenvolvida nesta pesquisa tem como foco a observação do objetivo geral do estudante ao longo da execução do MOOC.

2. MASSIVE OPEN ONLINE COURSES

Os Cursos Online Abertos e Massivos (MOOCs) constituem uma modalidade emergente de educação mediada por tecnologia caracterizada por oferecer cursos com acesso totalmente online, sem custo e sem requisitos formais de ingresso, permitindo que um grande número de participantes de diferentes contextos geográficos e sociais acesse conteúdos educacionais por meio de plataformas digitais. Essa forma de educação tem sido amplamente estudada nas últimas décadas, sobretudo em função de seu potencial de ampliar o acesso à educação formal e informal e de desafiar modelos tradicionais de ensino superior, apoiando práticas de aprendizagem ao longo da vida. A proliferação dos MOOCs se relaciona à crescente adoção de tecnologias educacionais digitais e à democratização do conhecimento em ambientes conectados digitalmente (Wang *et al.*, 2022).

Diferentes estudos destacam que, embora os MOOCs tenham apresentado expansão global significativa, o engajamento dos aprendizes e as taxas de conclusão ainda representam desafios centrais. Revisões sistemáticas da literatura evidenciam que grande parte dos participantes tende a abandonar os cursos antes de sua conclusão, o que tem sido interpretado tanto como um problema de aprendizagem quanto como reflexo das características flexíveis e abertas dessa modalidade. Nesse sentido, o padrão de engajamento dos estudantes — incluindo comportamento, cognição, dimensão emocional e interação social — é um dos principais fatores associados à persistência ou ao abandono no contexto dos MOOCs (Wang *et al.*, 2022).

A literatura também enfatiza que os MOOCs se distinguem de outras ofertas de ensino a distância por sua escala massiva e pela ausência de barreiras formais de ingresso, o que possibilita a participação de milhares ou mesmo milhões de estudantes simultaneamente, independentemente de sua afiliação institucional ou localização geográfica. Esses cursos frequentemente incluem recursos educacionais variados, como vídeos, fóruns de discussão, atividades avaliativas automatizadas e mecanismos de feedback, que juntos constituem um ambiente de aprendizagem interativo e flexível (Huang, Jew e Qi, 2023).

Apesar da promessa de democratização do acesso ao ensino, estudos recentes salientam que as taxas de evasão continuam sendo elevadas, sendo que muitos participantes interagem com os conteúdos de forma superficial ou não completam os requisitos formais de certificação. Essas tendências têm sido associadas não apenas a fatores instrucionais e de design do curso, mas também a questões que envolvem motivação intrínseca, competência digital e suporte pedagógico aos estudantes em contextos online (Wang *et al.*, 2022; Huang, Jew e Qi, 2023).

Nesse sentido, pesquisadores propõem que a compreensão profunda dos padrões de engajamento e dos fatores subjacentes à participação em MOOCs é essencial não apenas para melhorar práticas de design educacional, mas também para ampliar a eficácia dessas ofertas como instrumentos de aprendizagem. A literatura indica a necessidade de abordagens pedagógicas centradas no estudante e de estratégias que incentivem interações mais significativas, com potencial para reduzir as taxas de abandono e fortalecer os resultados educacionais.

3. MAPEAMENTO DE PERFIS DE ALUNOS

Na literatura, existe uma diversidade de padrões consolidados para a definição de perfis de alunos, dentre os quais se destacam os propostos por Kolb (1984), Honey e Mumford (1982) e Felder e

Silverman (1988). A proposta de Kolb (1984) fundamenta-se na teoria da aprendizagem experiencial, a qual descreve o processo de aprendizagem como um ciclo composto por quatro estágios e sustenta que os alunos necessitam de quatro habilidades para uma aprendizagem eficaz: experiência concreta, observação reflexiva, conceitualização abstrata e experimentação ativa. Essas habilidades organizam-se em duas dimensões polarizadas: Concreto/Abstrato e Ativo/Reflexivo.

De acordo com Kolb (1984), com base nessas dimensões, os alunos podem ser classificados em quatro perfis: Convergentes, que empregam a conceitualização abstrata e a experimentação ativa em seu processo de aprendizagem; Divergentes, que utilizam a vivência concreta e a observação reflexiva; Assimiladores, que se caracterizam pela observação reflexiva e pela conceitualização abstrata; e Acomodadores, que fazem uso da experimentação ativa e da vivência concreta.

O modelo proposto por Honey e Mumford (1982) também se baseia na teoria da aprendizagem experiencial e apresenta semelhanças com o modelo de Kolb (1984), classificando os alunos em quatro tipos: Ativista (correspondente aos Acomodadores), Teórico (Assimiladores), Pragmatista (Convergentes) e Reflexivo (Divergentes). Por fim, o modelo de Felder e Silverman (1988) considera quatro dimensões: Dedutivo/Intuitivo, Visual/Verbal, Ativo/Reflexivo e Sequenciais/Globais, as quais se baseiam, respectivamente, na forma como os alunos percebem, recebem, processam e compreendem as informações.

Embora essas padronizações estejam amplamente difundidas na literatura, considera-se que o desenvolvimento de um padrão de perfil específico para o contexto dos MOOC é mais adequado do que a simples aplicação de instrumentos de avaliação destinados a categorizar os alunos segundo modelos pré-existentes. Tal entendimento fundamenta-se em diferentes aspectos.

Primeiramente, os MOOC configuram um contexto particular de ensino e aprendizagem. Em segundo lugar, os usuários de plataformas on-line apresentam objetivos bastante distintos daqueles observados no sistema tradicional de ensino, incluindo, por exemplo, a obtenção de horas de atividades complementares, a necessidade de certificados para processos seletivos e/ou concursos, bem como a busca por certificações associadas a conhecimentos previamente adquiridos.

Por fim, a própria configuração dos cursos em plataformas digitais de aprendizagem apresenta grande variabilidade, com MOOC voltados a públicos específicos ou de ampla abrangência, com diferentes níveis de complexidade das atividades, distintos formatos de conteúdo (vídeos, textos ou ambos) e presença ou ausência de fóruns de discussão. Diante dessa diversidade, uma avaliação mais específica mostra-se pertinente para esse contexto.

Com base nesses aspectos, considerou-se mais adequada a adoção de um processo de mapeamento de perfis, entendido como uma representação capaz de capturar as principais características do aluno e fornecer o máximo de coerência em relação ao contexto estudado. Seu objetivo consiste em oferecer uma descrição eficiente dos diferentes aspectos do estudante — como comportamento, conhecimento e estilo de aprendizagem — não de forma individualizada, mas por meio do agrupamento de estudantes que compartilham atributos semelhantes (Hamim; Benabbou; Sael, 2019).

Para a realização do mapeamento de perfis de alunos, encontram-se disponíveis diversas abordagens. Entretanto, segundo Hamim, Benabbou e Sael (2019), as principais técnicas empregadas com esse propósito são:

1. Ontologias: estruturas de dados que representam conceitos ou grupos de conceitos, sendo úteis para o compartilhamento de conhecimento, a construção de consenso e o desenvolvimento de sistemas baseados em conhecimento.
2. E-portfólios (Portfólios Eletrônicos): conceito fundamentado na coleta de informações com o objetivo de rastrear o progresso do aluno ao longo de seu percurso educacional. Em geral, um portfólio eletrônico consiste em uma coleção de recursos digitais de aprendizagem e evidências que demonstram o desenvolvimento pessoal e as realizações do estudante.
3. Redes Bayesianas: modelos gráficos que permitem a representação de dependências entre eventos, normalmente simbolizados por meio de grafos.
4. *Machine Learning* (Aprendizagem de Máquina) e *Deep Learning* (Aprendizagem Profunda): técnicas amplamente empregadas em diferentes contextos e que também são utilizadas na identificação de padrões de perfis, especialmente quando se utilizam bases de dados como subsídio para o mapeamento, com o intuito de agrupar alunos em comunidades ou estereótipos que compartilham interesses comuns.

Além das técnicas mencionadas, Hamim, Benabbou e Sael (2019) estruturaram um conjunto de informações denominado Taxonomia de Características dos Alunos, que reúne os aspectos considerados fundamentais para a observação e análise no mapeamento de perfis comportamentais de alunos. A taxonomia proposta pelos autores encontra-se descrita no Quadro 1.

Quadro 1 – Descrição dos elementos da Taxonomia de Características para mapeamento de perfis de Alunos

Informações/Atributos Sobre os Alunos	Descrição
Identidade Pessoal	Define o aluno de uma forma única em comparação a outros, abrange: sexo, gênero, idade, nacionalidade, nome/sobrenome, profissão, código do aluno, telefone, e-mail.
Identidade Social	É um elemento derivado do pertencimento de um indivíduo a um grupo social, abrange: estado civil, educação dos pais, trabalho dos pais, religião, endereço, cultura.
Conhecimento	Refere-se às informações adquiridas por meio de experiências ou educação, pode ser: geral, teórica ou prática, em um tópico particular ou em geral.
Condições Físicas	Refere-se ao estado físico (habilidades) de um indivíduo (aluno), engloba elementos como: audição, visão e locomoção.
Aprendizagem	Estilo de Aprendizagem: Se refere a padrões de aprendizagem perceptíveis na conduta dos alunos, esses padrões podem ser enunciados a partir dos existentes na literatura como: Kolb (1984), Honey e Mumford (1982) e Felder; Silverman (1988), ou então mapeados particularmente para o contexto estudado.
	Objetivos de Aprendizagem: Individual de cada aluno, como: aprender um tópico específico, conseguir um certificado ou certificar um conhecimento pré-existente.
	Comportamento de Aprendizagem: <i>Comportamento geral</i> – Atenção, tomada de decisão, responsabilidade, autonomia, motivação (autoestima, satisfação no trabalho, interesse intrínseco, auto eficácia); <i>Comportamento on-line</i> – comentários e postagens, participação em enquetes, realização de testes, sequência de navegação, acesso a recursos (frequência, duração, data), equipamento utilizado (tablet, computador, celular); <i>Comportamento presencial:</i> participação nas aulas, tempo de estudo em casa, atividades pedagógicas, compreensão dos conteúdos.
Cognitivo	Refere-se a uma série de processos mentais relacionados a manipulação da informação, inclui: inteligência, competência, experiência, habilidades, aspectos metacognitivos, que se referem às cognições do aluno sobre a sua própria cognição, autoavaliação, autorregulação, autogestão.

Acadêmico	Diz respeito às informações relacionadas ao histórico acadêmico do aluno, diplomas, como de especialização, bacharelado, ensino médio ou ensino fundamental.
Psicológico	Diz respeito à saúde mental do aluno e a sua personalidade, abrange: depressão, vícios, ansiedade, estado de espírito, transtornos alimentares, personalidade (introverso, extroverso, intuitivo).
Habilidade e Interesses	Refere-se aos talentos em função dos interesses dos alunos, talentos que permitam alcançar objetivos particulares, abrangendo: criatividade, interpessoal, comunicação, liderança, compreensão (velocidade e lógica).
Linguagem	Idioma preferido, nível de idiomas falados e escritos.

Fonte: Adaptado de Hamim, Benabbou e Sael (2019).

Para este estudo foi utilizado como base para a montagem da base de dados a Taxonomia de Características dos Alunos (Hamim, Benabbou e Sael, 2019), e para realização do mapeamento dos perfis foi empregada a técnica de *Machine Learning*.

4. TRABALHOS RELACIONADOS: ESTUDOS CLÁSSICOS E ABORDAGENS CONTEMPORÂNEAS

Nos últimos anos, e especialmente desde a consolidação dos Massive Open Online Courses, diversos estudos vêm sendo desenvolvidos com o objetivo de mapear perfis e comportamentos de alunos em ambientes de aprendizagem on-line. Essas investigações evidenciam que a preocupação com a compreensão do comportamento discente em MOOC não é recente, estendendo-se desde os primeiros anos de popularização desse formato educacional até pesquisas contemporâneas que incorporam técnicas avançadas de análise de dados. Entre os trabalhos que fundamentam essa trajetória investigativa destacam-se estudos clássicos, como os de Rodrigues *et al.* (2016), Tan *et al.* (2018), Korchi, Abdellah e Fez (2015), Rezgui, Mhiri e Ghédira (2017), Ferreira *et al.* (2017), Finney *et al.* (2018) e Snyman e Van Den Berg (2018), bem como pesquisas mais recentes que ampliam e aprofundam essa discussão, a exemplo de Chen *et al.* (2025), Silva *et al.* (2024), Cagiltay, Toker e Cagiltay (2024), Luna *et al.* (2024) e El Ayady *et al.* (2025).

Rodrigues *et al.* (2016) desenvolveram uma pesquisa no contexto de MOOC com o objetivo de identificar perfis de engajamento dos alunos. Para tanto, analisaram dados de aproximadamente 5 mil estudantes do curso “Nova Gramática da Língua Portuguesa”, ofertado pela plataforma Openredu, considerando atributos relacionados à frequência de postagens, assiduidade e desempenho. A aplicação do algoritmo de agrupamento K-means possibilitou a identificação de três perfis de engajamento — Engajados, Esporádicos e Desengajados — evidenciando o potencial das técnicas de agrupamento para a identificação de padrões comportamentais em MOOC.

Tan *et al.* (2018) também empregaram o algoritmo K-means para mapear perfis de aprendizagem de alunos no MOOC “E-learning e Culturas Digitais”, desenvolvido pela Universidade de Edimburgo. O estudo considerou atributos comportamentais, dados autorrelatados e resultados de avaliações, culminando na identificação de quatro perfis distintos de aprendizagem. Os autores destacam que o reconhecimento desses perfis pode subsidiar estratégias de personalização e aprimoramento do design instrucional em MOOC.

No que se refere a abordagens baseadas em Ontologias, Korchi, Abdellah e Fez (2015) propuseram um modelo conceitual para o reconhecimento de perfis de alunos em ambientes de aprendizagem assistidos por computador, estruturado a partir de múltiplas dimensões de dados. De modo

complementar, Ferreira *et al.* (2017) apresentaram uma abordagem híbrida que integra Ontologias e Redes Bayesianas para mapear o comportamento discente em Ambientes Virtuais de Aprendizagem, permitindo inferências mais consistentes ao considerar as incertezas inerentes ao processo de aprendizagem.

Outras abordagens clássicas incluem o uso de Portfólios Eletrônicos, como proposto por Rezgui, Mhiri e Ghédira (2017), que apresentaram um modelo conceitual para documentar o progresso e as competências dos alunos ao longo do tempo, e estudos empíricos como o de Finney *et al.* (2018), que investigaram perfis de busca de ajuda em estudantes universitários, e o de Snyman e Van Den Berg (2018), que analisaram narrativas e histórias de vida para compreender trajetórias de aprendizagem e desenvolvimento pessoal.

Mais recentemente, as pesquisas passaram a incorporar técnicas mais avançadas e análises em larga escala. Chen *et al.* (2025) propuseram um modelo de perfil de usuários de MOOC baseado nas dimensões de emoção, comportamento e habilidade, utilizando técnicas de agrupamento aplicadas a grandes volumes de dados educacionais. Esse estudo amplia a compreensão dos perfis de aprendizagem ao integrar indicadores emocionais às métricas tradicionais de engajamento.

No contexto brasileiro, Silva *et al.* (2024) investigaram o engajamento e as influências demográficas em MOOCs ofertados pela plataforma Aprenda Mais, evidenciando que variáveis como raça/cor, deficiência e localização geográfica impactam significativamente os padrões de participação e conclusão dos cursos, reforçando a necessidade de estratégias educacionais mais inclusivas.

Cagiltay, Toker e Cagiltay (2024) analisaram padrões comportamentais de estudantes em MOOC considerando variáveis demográficas e número de matrículas, revelando tendências relevantes de participação e fornecendo subsídios para melhorias em iniciativas educacionais de larga escala. Já Luna *et al.* (2024) aplicaram técnicas de descoberta de subgrupos para caracterizar diferentes tipos de aprendizes em MOOC, demonstrando o potencial dessa abordagem para a personalização de conteúdos e recomendações educacionais.

Por fim, El Ayady *et al.* (2025) exploraram a previsão de risco de insucesso acadêmico em MOOC por meio da análise de séries temporais multivariadas de interações comportamentais, evidenciando a possibilidade de antecipar dificuldades de aprendizagem e implementar intervenções pedagógicas mais eficazes.

De forma geral, a literatura evidencia que o mapeamento de perfis de alunos em MOOC constitui uma linha de pesquisa consolidada e em constante evolução, na qual abordagens clássicas fornecem bases conceituais e metodológicas que vêm sendo ampliadas por estudos contemporâneos. Essa continuidade reforça a relevância do presente estudo, que se insere nesse contexto ao propor o mapeamento de perfis comportamentais de alunos a partir da análise de dados educacionais, contribuindo para o avanço das estratégias de personalização e melhoria dos ambientes de aprendizagem on-line.

5. METODOLOGIA

Após a investigação da literatura e a reflexão acerca das diferentes técnicas passíveis de aplicação, foi desenvolvida uma abordagem metodológica para o mapeamento dos perfis comportamentais dos estudantes usuários do MOOC investigado. Tal abordagem fundamenta-se na técnica de

Agrupamento (Clustering), vinculada à Aprendizagem de Máquina, bem como na análise e exploração dos grupos resultantes do processo de agrupamento.

No que se refere à Aprendizagem de Máquina, foi empregado o algoritmo de agrupamento K-means, amplamente utilizado em estudos correlatos, como os de Rodrigues *et al.* (2016) e Tan *et al.* (2018). A aplicação desse algoritmo demandou a construção e formatação prévia de uma base de dados estruturada contendo informações dos alunos, de modo a viabilizar a identificação de padrões comportamentais a partir dos dados disponíveis.

Para a definição dos atributos que compuseram essa base de dados, foram considerados os estudos de Korchi, Abdellah e Fez (2015), Ferreira *et al.* (2017) e Hamim, Benabbou e Sael (2019), os quais forneceram subsídios teóricos e metodológicos para a compreensão dos principais aspectos a serem observados no mapeamento de perfis de alunos em ambientes educacionais on-line. Nesse contexto, adotou-se, de forma particular, a Taxonomia de Características dos Alunos proposta por Hamim, Benabbou e Sael (2019), na qual os autores sistematizam, de maneira categórica e concisa, os atributos considerados fundamentais para a caracterização comportamental dos estudantes.

Após a execução do algoritmo K-means e a consequente geração dos agrupamentos, realizou-se uma investigação analítica com o objetivo de examinar as características predominantes de cada grupo identificado. A partir dessa análise, foi possível interpretar os padrões comportamentais associados aos agrupamentos e, consequentemente, definir os perfis dos alunos que cursaram o MOOC analisado.

5.1 Descrição do MOOC e Participantes

O MOOC investigado pertence à área de Química e é ofertado por uma plataforma brasileira de cursos abertos on-line. O curso possui carga horária total de 15 horas e caracteriza-se por apresentar um público bastante diversificado, uma vez que não exige pré-requisitos para participação, não impõe restrições para o cadastramento dos interessados e é disponibilizado de forma gratuita.

A estrutura do curso é composta por quatro módulos de conteúdos e um módulo inicial introdutório. Esse módulo inicial apresenta orientações sobre o funcionamento do MOOC, disponibiliza materiais introdutórios e inclui um questionário destinado ao levantamento do perfil dos estudantes. O principal recurso educacional utilizado ao longo do curso consiste em videoaulas gravadas pelo professor responsável, complementadas por materiais em formato textual. Todos os módulos de conteúdo possuem uma atividade avaliativa, estruturada no formato de questionários de múltipla escolha. Ao final do curso, os alunos têm acesso a um fórum de discussão, que possibilita a troca de informações, a interação entre os participantes e a manifestação de opiniões sobre o curso. Além disso, é disponibilizado um material complementar elaborado pelo docente e, após a conclusão, o aluno pode emitir o certificado correspondente.

Até o primeiro semestre de 2025, o MOOC contava com mais de 5.000 alunos matriculados. Para a realização deste estudo, foram utilizados dados de 3.540 estudantes que responderam ao questionário de perfil do aluno até a data de extração das bases de dados. No que se refere ao abandono do curso, não foi possível identificar de forma precisa quais alunos haviam desistido ou quais ainda estavam em processo de realização do MOOC, uma vez que o curso não possui prazo limite para conclusão. Além disso, a base de dados contempla estudantes que demandaram períodos distintos para concluir o curso e emitir o certificado, variando de forma significativa entre os participantes.

Diante desse contexto, mesmo considerando informações como o tempo médio de permanência no curso, não foi possível afirmar com exatidão quais estudantes evadiram. Assim, para fins deste estudo, foram considerados como concluintes os alunos que emitiram o certificado até a data de extração dos dados, e como não concluintes aqueles que não realizaram a emissão do certificado até esse momento.

5.2 Atributos e Coleta de Dados

Para o desenvolvimento deste estudo, foram extraídos dados provenientes de quatro relatórios disponibilizados pela plataforma MOOC analisada, a saber: Perfil do Estudante, Conclusão de Atividades, Notas e Logs. De cada um desses relatórios, foram selecionadas informações consideradas relevantes para os objetivos da pesquisa, as quais foram posteriormente integradas em uma única base de dados consolidada.

Os dados que compõem essa base foram obtidos a partir das interações dos estudantes ocorridas no período compreendido entre junho de 2019 e agosto de 2020. Sobre esse conjunto de dados, foram realizados procedimentos de transformação e pré-processamento, com o intuito de aprimorar a qualidade das informações e torná-las mais adequadas à análise dos perfis comportamentais dos alunos. Os atributos que compõem a base de dados final, utilizada para a análise e para a definição dos perfis dos estudantes, são apresentados no Quadro 2, no qual se descrevem as variáveis consideradas no processo de mapeamento.

Quadro 2 – Atributos da Base de Dados

ID	Atributo	Definição	Relatório de Origem
1	Gênero	Informado pelo usuário.	Extraído do Relatório de Perfil do Estudante.
2	Idade	Informada pelo usuário.	Extraído do Relatório de Perfil do Estudante.
3	Escolaridade	Informada pelo usuário – vai desde o fundamental incompleto até a pós-graduação.	Extraído do Relatório de Perfil do Estudante.
4	Motivação	Informada pelo usuário – que pode ser: curiosidade geral, importante para meu trabalho, ou importante para meus estudos.	Extraído do Relatório de Perfil do Estudante.
5	Intenção	Informada pelo usuário – diz respeito a vontade de terminar o MOOC, pode ser: Sim, Não sei, ou Não.	Extraído do Relatório de Perfil do Estudante.
6	Nota Final	É a média das notas dos 4 questionários realizados.	Extraído do Relatório de Notas.
7	Tempo	É a diferença de tempo, dada em dias, entre o primeiro acesso a alguma atividade do MOOC até o último acesso – tempo que o aluno interagiu com o curso.	Calculado a partir do Relatório de Conclusão de Atividades.
8	Porcentagem Concluída	É a porcentagem das atividades concluídas – exemplo se no curso há 20 recursos e/ou atividades e o aluno concluiu 10 ele tem 0,5 de porcentagem concluída.	Calculado a partir do Relatório de Conclusão de Atividades.
9	Soma das Atividades	Nesse atributo foram somados quantas vezes os alunos acessaram os	Calculado a partir do Relatório de Logs.

	Realizadas – Materiais em Texto	materiais em texto. É o somatório apenas as atividades referentes aos conteúdos do MOOC.	
10	Soma das Atividades Realizadas – Vídeo Aulas	Nesse atributo foram somados quantas vezes os alunos acessaram as vídeo aulas.	Calculado a partir do Relatório de Logs.
11	Soma das Atividades Realizadas – Tentativas em Questionários Avaliativos	Nesse atributo foram somados quantas vezes os alunos acessaram as tentativas de questionário.	Calculado a partir do Relatório de Logs.
13	Soma das Atividades Realizadas	Nesse atributo foram somados todos os acessos a recursos de conteúdos e atividades avaliativas do MOOC. Também definido como a quantidade de interações com a plataforma.	Calculado a partir do Relatório de Logs.
13	Certificado	Variável binária que tem valor 1 para quem conclui e gerou o certificado e 0 para quem não requisitou o certificado.	Extraído do Relatório de Conclusão de Atividades.

Fonte: Autora.

5.3 Aplicação do algoritmo de Agrupamento K-means

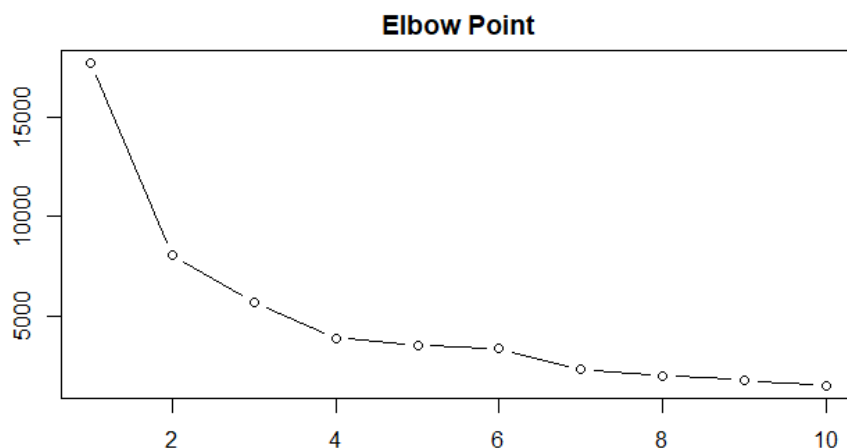
Após a formatação da base de dados para a análise dos perfis comportamentais, foram selecionados apenas os atributos considerados mais relevantes para a aplicação do algoritmo de agrupamento K-means. Essa seleção levou em conta, sobretudo, as características do próprio algoritmo, que utiliza a distância entre dois pontos — especificamente a distância Euclidiana — para a formação dos agrupamentos por similaridade. Em razão disso, não é possível a utilização de atributos categóricos, sendo considerados apenas atributos numéricos.

Dessa forma, a base de dados submetida ao algoritmo foi composta pelos seguintes atributos: Tempo, Porcentagem Concluída, Tentativas em Questionários Avaliativos, Soma Total de Atividades Realizadas e Certificado. A aplicação do algoritmo K-means foi realizada por meio da linguagem de programação e ciência de dados R.

Antes da submissão da base de dados ao algoritmo, foi realizado o teste conhecido como *elbow method* (tradução livre: método do cotovelo), com o objetivo de identificar a quantidade de agrupamentos que melhor se ajustaria ao conjunto de dados. O *elbow method* é uma heurística amplamente utilizada para a determinação do número ideal de clusters, consistindo na plotagem da variação explicada em função do número de grupos, sendo o ponto de inflexão da curva — o “cotovelo” — considerado o número mais adequado de agrupamentos.

No caso da base de dados utilizada neste estudo, os valores indicados situaram-se entre 2 e 4 agrupamentos, conforme ilustrado na Figura 2. Observa-se, entretanto, que o ponto de cotovelo ocorre em $k = 2$, valor que, segundo o método, representa a melhor escolha para a configuração dos agrupamentos.

Figura 2 – Método Elbow Point



Fonte: Autor.

Entretanto, optou-se por realizar testes iniciais considerando quatro agrupamentos, levando-se em conta o conhecimento prévio sobre a base de dados analisada. Contudo, observou-se que, nessa configuração, um dos grupos formados apresentava um número muito reduzido de participantes (apenas 47 alunos), não caracterizando um perfil comportamental singular em relação aos objetivos dos estudantes no curso, diferentemente dos demais grupos identificados.

Diante disso, definiu-se a quantidade de agrupamentos como três. Com essa configuração, a aplicação do algoritmo K-means possibilitou o agrupamento de alunos que apresentaram atributos semelhantes, resultando em perfis mais consistentes e interpretáveis. Foram utilizadas as configurações *default* do algoritmo, nas quais os dados são agrupados pelo método das k-médias, cujo objetivo é particionar os pontos em k grupos de forma a minimizar a soma dos quadrados das distâncias entre os pontos e os centros dos respectivos agrupamentos. Na linguagem R, o algoritmo de Hartigan e Wong (1979) é empregado por padrão para a execução do K-means.

Cabe destacar que, para a aplicação do algoritmo K-means — uma vez que este utiliza a distância entre pontos como métrica de similaridade — é fundamental realizar a padronização das escalas dos atributos. Neste estudo, esse procedimento foi realizado por meio da normalização dos dados.

Assim, o algoritmo K-means foi configurado com $k = 3$ e aplicado à base de dados. Em seguida, foi conduzida uma investigação das características dos atributos de cada grupo identificado, com a sistematização dos dados de perfil e ênfase nas ações realizadas pelos estudantes na plataforma. Foram analisados, em especial, o número médio de interações dos grupos, o tempo de realização do curso, a porcentagem de atividades concluídas e o desempenho acadêmico. Ressalta-se que este estudo tem como propósito mapear os perfis dos alunos com base em seus objetivos ao realizarem um MOOC, a partir da análise de seus comportamentos e padrões de uso da plataforma.

6. RESULTADOS

Nesta seção são apresentadas e interpretadas as análises realizadas a partir da base de dados do estudo. Inicialmente, descrevem-se as características gerais da amostra e dos agrupamentos gerados pelo algoritmo de agrupamento. Em seguida, com base na análise dos atributos e dos padrões identificados, são definidos e caracterizados os perfis comportamentais dos estudantes participantes do MOOC.

6.1 Análise Descritiva dos Dados

Os estudantes participantes do curso, conforme já mencionado, compõem um público heterogêneo. Observa-se a presença de alunos com idades variando entre 19 e 50 anos, com predominância daqueles com até 19 anos, que representam 38% da amostra analisada. No que se refere ao nível de escolaridade, os participantes apresentam formações que variam do Ensino Fundamental Incompleto à Pós-Graduação, destacando-se a maior concentração de alunos com Ensino Superior Incompleto, correspondendo a 36% dos estudantes do MOOC.

Quanto à motivação para participação no curso, 46% dos alunos indicaram que o interesse está relacionado à importância do conteúdo para seus estudos. Além disso, a grande maioria dos participantes (96%) declarou intenção de concluir o MOOC. Entretanto, apenas 24% dos estudantes efetivamente finalizaram o curso, evidenciando uma diferença significativa entre a intenção declarada e a conclusão efetiva.

Em relação ao desempenho acadêmico, verificou-se que 35% dos alunos obtiveram notas finais entre 30 e 40 pontos. Cabe ressaltar que a avaliação do MOOC é composta por quatro questionários avaliativos, cada um com valor máximo de 10 pontos, sendo a nota final calculada a partir do somatório das pontuações obtidas em cada avaliação. Informações mais detalhadas sobre o perfil da amostra analisada são apresentadas no Quadro 3.

Quadro 3 – Descrição da Amostra

Atributo		Quantidade absoluta de Alunos
Gênero	Masculino	2.074
	Feminino	1.457
	Outros	9
Idade	Até 19 anos	1.375
	20 a 24 anos	903
	25 a 29 anos	417
	30 a 34 anos	242
	35 a 39 anos	209
	40 a 49 anos	244
	Acima de 50 anos	150
Escolaridade	Ensino fundamental incompleto	385
	Ensino fundamental completo	69
	Ensino médio incompleto	438
	Ensino médio completo	373
	Ensino superior incompleto	1.292
	Ensino superior completo	435
	Pós-Graduação	548
Motivação	Curiosidade geral	1.474
	Importante para meu trabalho	421
	Importante para meus estudos	1.645
Intenção	Não	6
	Não Sei	135
	Sim	3.399
Nota Final	0	1.602
	1-10	387

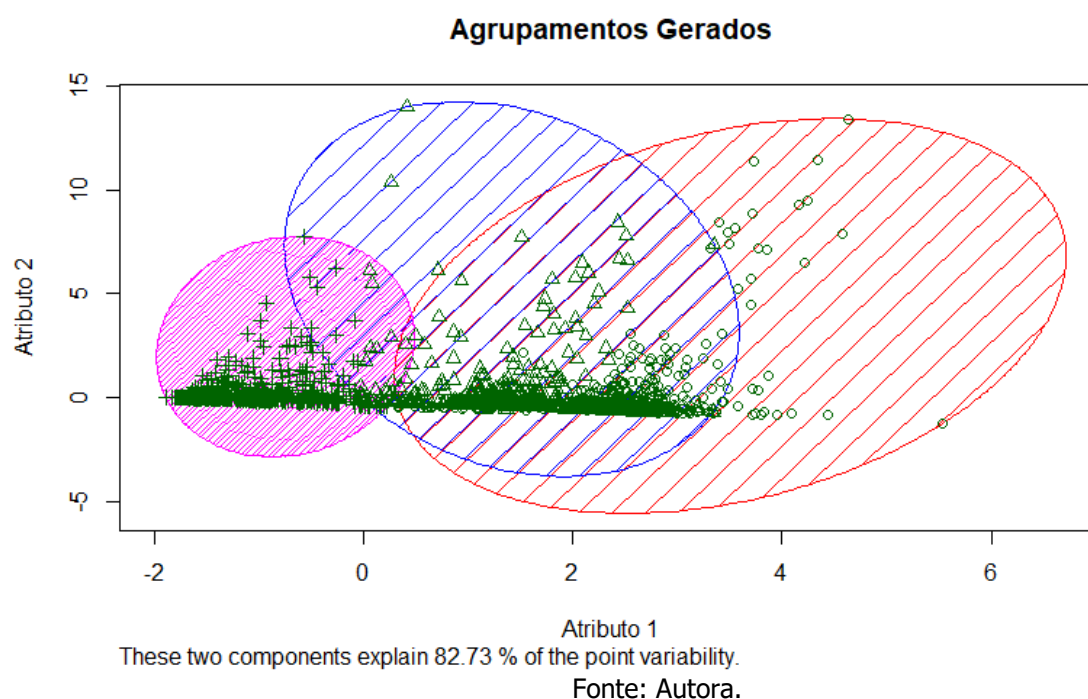
	11-20	116
	21-30	181
	31-40	1.254
Certificado	Concluintes	876
	Não Concluintes	2.664
TOTAL DE ALUNOS		3.540

Fonte: Autora.

6.2 Análise Descritiva dos Agrupamentos

Com a aplicação do algoritmo K-means, os estudantes foram distribuídos em três agrupamentos distintos. Observa-se uma relativa sobreposição entre alguns atributos dos alunos, o que é esperado em processos de agrupamento baseados em similaridade. Ainda assim, a visualização gráfica dos agrupamentos (Figura 2) permite compreender de forma mais clara a distribuição dos grupos identificados.

Figura 2 – Gráfico dos Agrupamentos Gerados.



Os atributos utilizados para a construção do gráfico dos agrupamentos foram selecionados automaticamente pela biblioteca cluster, por apresentarem maior influência na formação dos grupos. Esses atributos são capazes de explicar 82,73% da variância total dos dados, indicando que a representação gráfica sintetiza de maneira adequada as principais diferenças entre os agrupamentos gerados pelo algoritmo.

1. **Agrupamento 1 (vermelho):** O Agrupamento 1 é composto por 862 estudantes, dos quais 51% são do sexo masculino. A faixa etária dos participantes varia entre 19 e 50 anos, com predominância de estudantes mais jovens, destacando-se a faixa de 20 a 24 anos, que representa 36% do grupo. Em relação à escolaridade, observa-se que 49% dos alunos possuem ensino superior incompleto. No que se refere à motivação para a realização do curso, 44% dos participantes afirmaram ter se inscrito por curiosidade geral, outros 44% relataram que o tema é importante para seus estudos, e 12% indicaram relevância para o trabalho.

Quanto à intenção de conclusão, 95% dos estudantes declararam pretender finalizar o curso e, efetivamente, 100% dos alunos deste agrupamento concluíram o MOOC, com emissão do certificado.

Em termos de interação com a plataforma, esse grupo apresentou elevada participação, com média de 108 interações por estudante. Observou-se, em média, três visualizações por videoaula, 11 acessos a cada questionário avaliativo e três acessos a cada material em formato textual. Os estudantes concluíram, em média, 96% das atividades propostas e permaneceram aproximadamente 13 dias no curso, considerando o intervalo entre o primeiro acesso e a emissão do certificado. Quanto ao desempenho, a nota final média foi de 36,5 pontos, o que corresponde a aproximadamente 9,12 pontos por questionário avaliativo.

2. **Agrupamento 2 (azul):** O Agrupamento 2 é constituído por 583 alunos, sendo 53% do sexo masculino. A maior concentração etária encontra-se entre estudantes com até 19 anos, que representam 36% do grupo, seguidos pela faixa de 20 a 24 anos, correspondente a 20% dos participantes. Em relação à escolaridade, predomina o ensino superior incompleto, relatado por 38% dos alunos, seguido pela pós-graduação, com 20%. Quanto à motivação, 51% dos estudantes indicaram que o principal fator para a matrícula foi o fato de o tema ser importante para seus estudos. Assim como no agrupamento anterior, 95% dos participantes afirmaram ter intenção de concluir o curso.

Apesar dessa elevada intenção declarada, nenhum dos alunos pertencentes a esse agrupamento concluiu o MOOC, resultando em uma taxa de não conclusão de 100%. Ainda assim, observa-se um nível elevado de interação com a plataforma, semelhante ao do Agrupamento 1, com média de 108 interações por aluno. Os estudantes realizaram, em média, três acessos a cada videoaula e a cada material textual, além de aproximadamente 11 acessos a cada questionário avaliativo. A porcentagem média de conclusão das atividades foi de 87%, e o tempo médio de permanência no curso foi de 16 dias. O desempenho médio dos alunos desse grupo foi de 33,4 pontos na nota final, equivalente a cerca de 8,35 pontos por questionário.

3. **Agrupamento 3 (roxo):** O Agrupamento 3 é o mais numeroso, reunindo 2.095 estudantes, dos quais 66% são do sexo feminino. A maioria dos participantes (42%) encontra-se na faixa etária de até 19 anos, embora estejam representadas todas as demais faixas etárias. No que diz respeito à escolaridade, 30% dos alunos relataram possuir ensino superior incompleto. Quanto à motivação, 46% indicaram que o principal motivo para a realização do curso foi a importância do tema para seus estudos, e 96% declararam intenção de concluir o MOOC.

Apesar da elevada intenção declarada, esse agrupamento apresentou uma taxa de não conclusão bastante elevada, com 99% dos alunos não emitindo o certificado. No que se refere às interações com a plataforma, observam-se médias significativamente inferiores às dos demais grupos, com aproximadamente 21 interações por estudante ao longo de todo o curso. Considerando que o MOOC é composto por 24 atividades, incluindo videoaulas, materiais textuais e questionários avaliativos, esse número indica que, em muitos casos, os alunos não acessaram sequer uma vez cada recurso disponibilizado. Consequentemente, a média de conclusão das atividades foi de aproximadamente 8%, o tempo médio de permanência no curso foi de apenas dois dias, e a nota final média obtida pelos estudantes ficou em torno de 2,8 pontos.

Mais detalhes referentes aos valores dos atributos dos agrupamentos identificados podem ser observados no Quadro 4.

Quadro 4 – Médias e Medianas dos atributos por grupos

Atributo	Métrica	Geral	Grupo 1	Grupo 2	Grupo 3
Tempo	Mediana	0,02	2,4	1,1	0,0
	Média	7,4	13,2	16,3	2,6
Porcentagem Concluída	Mediana	19 %	100,0 %	96 %	3,8 %
	Média	42 %	96,9 %	86 %	8,5 %
Soma dos Acessos aos Materiais em Texto	Mediana	7	12,0	12,0	0,0
	Média	6,76	13,2	13,1	2,3
Soma dos Acessos às Vídeo Aulas	Mediana	25	50,0	50,0	5,0
	Média	26,31	51,5	51,3	8,9
Soma dos Acessos aos Questionários	Mediana	26	43,5	40,0	0,0
	Média	23,79	43,9	44,4	9,7
Soma das Atividades Realizadas	Mediana	60	108,0	104,0	7,0
	Média	56,86	108,7	108,9	21,1

Fonte: Autora.

6.3 Definição dos Perfis Comportamentais

A partir da geração dos agrupamentos por meio do algoritmo K-means, foi possível compreender por que o *elbow method* indicou inicialmente que a melhor divisão da base de dados ocorreria em dois agrupamentos, uma vez que tal método se apoia, de forma predominante, na distinção entre alunos concluintes e não concluintes. Entretanto, essa divisão binária revela-se excessivamente simplificada para a compreensão dos diferentes objetivos e comportamentos dos estudantes em MOOC. Diante disso, optou-se pela geração de três agrupamentos, tanto em função da familiaridade dos autores com a base de dados quanto em razão dos testes exploratórios realizados com diferentes valores de k.

A partir da análise detalhada dos três agrupamentos descritos anteriormente, foi possível identificar quatro perfis comportamentais distintos, definidos com base nas características dos grupos gerados pelo algoritmo e considerando, sobretudo, os objetivos predominantes dos estudantes ao realizarem o MOOC analisado.

- **Engajados – correspondem majoritariamente aos estudantes do Agrupamento 1.** O principal indicativo do engajamento desses alunos é a elevada quantidade de interações com a plataforma, com média de aproximadamente 108 interações ao longo do curso. Observa-se, ainda, que o tempo médio de permanência no MOOC, em torno de 13 dias, é

compatível com esse nível de interação. Esses estudantes concluíram, em média, cerca de 96% das atividades propostas, evidenciando amplo acesso aos recursos disponibilizados. O desempenho acadêmico reforça esse comportamento, uma vez que as notas médias por questionário se mantiveram acima de 9 pontos. Outro fator relevante é a distribuição equilibrada das interações entre os diferentes tipos de recursos, com média de acesso igual ou superior a três por tipo de material. Além disso, todos os alunos desse perfil emitiram o certificado de conclusão. Dessa forma, infere-se que o objetivo predominante dos estudantes engajados é a aprendizagem integral do conteúdo do curso, aliada à obtenção da certificação.

- **Estratégicos – correspondem majoritariamente aos estudantes do Agrupamento 2.** Embora nenhum dos alunos desse agrupamento tenha concluído o MOOC com a emissão do certificado, observa-se um padrão de comportamento bastante semelhante ao dos estudantes engajados. A média de interações com a plataforma também foi de aproximadamente 108, com cerca de três acessos a cada videoaula e a cada material textual. A porcentagem média de conclusão das atividades foi de aproximadamente 87%, e o tempo médio de permanência no curso foi de cerca de 16 dias. O desempenho acadêmico também se manteve elevado, com médias superiores a 8 pontos por questionário. A partir desses indicadores, infere-se que os estudantes com perfil estratégico possuíam como principal objetivo a aprendizagem de conteúdos específicos ou de determinados tópicos do curso, interrompendo a participação após atingirem esse objetivo, sem necessariamente buscar a certificação. Ressalta-se, ainda, que parte desses alunos pode ter emitido o certificado após o período de extração dos dados.

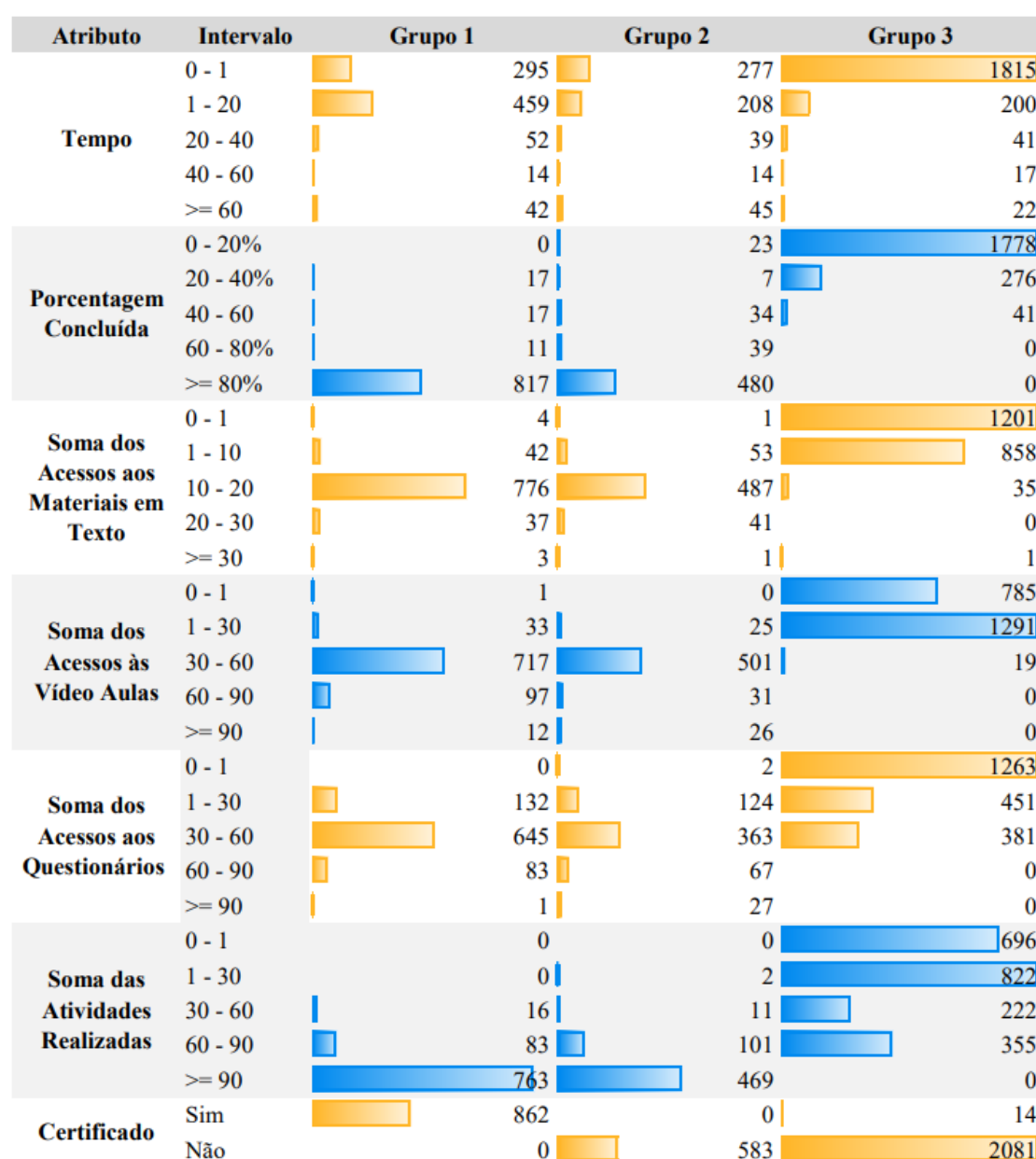
Devido, sobretudo, à elevada quantidade de estudantes no Agrupamento 3 (2.095 alunos), foi possível identificar dois perfis comportamentais distintos nesse grupo, considerando os padrões de interação e conclusão do curso.

- **Inativos – correspondem aos estudantes do Agrupamento 3 que não concluíram o MOOC.** Esse perfil caracteriza-se por níveis muito baixos de interação com a plataforma, com média de aproximadamente 21 interações por aluno. A porcentagem média de atividades concluídas foi de cerca de 8%, e o tempo médio de permanência no curso foi de apenas dois dias. Esses indicadores sugerem que os estudantes desse perfil acessaram apenas os primeiros recursos do MOOC e, em seguida, abandonaram o curso. Tal comportamento pode estar associado a diferentes fatores, como expectativas não atendidas em relação ao conteúdo ou ao formato do curso.
- **Instrumentais – correspondem aos estudantes do Agrupamento 3 que concluíram o curso.** Embora esse perfil seja predominante entre os alunos concluintes desse agrupamento, também foi possível identificar, em menor frequência, estudantes com características semelhantes nos Agrupamentos 1 e 2. Os alunos instrumentais apresentam um comportamento geral de baixa interação com a plataforma, mas ainda assim conseguem concluir o curso. Em média, realizaram cerca de 33 interações, das quais aproximadamente 31 concentram-se em tentativas de questionários avaliativos, restando apenas cerca de duas interações distribuídas entre os demais materiais do curso, que totalizam 20 recursos. A porcentagem média de conclusão das atividades foi relativamente baixa, em torno de 22%. Um aspecto particularmente relevante é o tempo médio de conclusão do curso, inferior a um dia, indicando que esses estudantes realizaram o primeiro acesso e emitiram o certificado no mesmo dia. O elevado número de tentativas nos questionários, aliado à baixa exploração dos demais materiais, sugere que o principal objetivo desses alunos é a obtenção da certificação,

utilizando o curso de forma estritamente instrumental. Além disso, especula-se que esses estudantes possuam pouco conhecimento prévio sobre o conteúdo, uma vez que necessitam de múltiplas tentativas para alcançar a pontuação mínima exigida nas avaliações.

As características do perfil instrumental podem ser visualizadas de forma mais ilustrativa na Figura 3, que apresenta a distribuição dos estudantes nos agrupamentos identificados pelo K-means. Observa-se, por exemplo, que no Agrupamento 1, apesar de todos os alunos terem emitido o certificado, 295 estudantes concluíram o curso em apenas um dia. Já no Agrupamento 3, nota-se que a maioria dos estudantes apresenta baixos índices de atividades concluídas, tempo de permanência e interações com a plataforma, configurando um comportamento típico de abandono. Entretanto, uma parcela desses estudantes conclui o curso mesmo com níveis mínimos de interação, característica distintiva do perfil instrumental.

Figura 3: Distribuição dos estudantes por Agrupamento



Fonte: Autora.

7. CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo mapear perfis comportamentais de estudantes participantes de um MOOC ofertado por uma plataforma brasileira, com ênfase na compreensão dos objetivos e interesses que orientam a participação desses alunos ao longo do curso. Para tal, foi proposta e aplicada uma abordagem baseada em técnicas de Aprendizagem de Máquina, especificamente o algoritmo de agrupamento K-means, aliada à análise interpretativa dos grupos gerados a partir de atributos comportamentais extraídos da plataforma.

Os resultados obtidos evidenciam que o comportamento dos estudantes em MOOC não pode ser adequadamente compreendido apenas pela dicotomia entre concluintes e não concluintes, como apontado implicitamente pelo *elbow method*. Em consonância com estudos anteriores (Rodrigues *et al.*, 2016; Tan *et al.*, 2018), observou-se que alunos com níveis semelhantes de interação e desempenho podem apresentar objetivos distintos ao longo do curso. A análise aprofundada dos agrupamentos permitiu identificar quatro perfis comportamentais predominantes: (1) Engajados, estudantes que interagem amplamente com os recursos do curso, concluem a maioria das atividades e buscam tanto a aprendizagem quanto a certificação; (2) Estratégicos, alunos que demonstram elevado envolvimento com conteúdos específicos, mas não necessariamente com a totalidade do curso ou com a obtenção do certificado; (3) Inativos, caracterizados por níveis mínimos de interação e elevada taxa de abandono precoce; e (4) Instrumentais, estudantes cujo comportamento é orientado prioritariamente à obtenção do certificado, com uso seletivo e restrito dos recursos educacionais, concentrando suas interações nas atividades avaliativas.

A identificação desses perfis dialoga diretamente com achados da literatura recente, que apontam a diversidade de motivações e estratégias de participação em MOOCs como um fenômeno recorrente e persistente ao longo do tempo (Korchi; Abdellah; Fez, 2015; Ferreira *et al.*, 2017; Hamim; Benabbou; Sael, 2019; Chen *et al.*, 2025). Em particular, o perfil instrumental reforça discussões contemporâneas sobre o uso pragmático dos MOOC, nos quais a certificação pode assumir maior relevância do que o processo formativo em si, aspecto também observado em investigações mais recentes sobre comportamento e desempenho em ambientes massivos de aprendizagem on-line.

Do ponto de vista prático, os resultados deste estudo oferecem subsídios relevantes para o aprimoramento do desenho pedagógico e da gestão de MOOC. A partir da abordagem proposta, torna-se possível aplicar novos testes em outros cursos da plataforma, com o intuito de validar a recorrência desses perfis e verificar a possibilidade de generalização dos resultados para diferentes áreas do conhecimento e formatos de curso. A consolidação empírica desses perfis pode apoiar a implementação de estratégias diferenciadas de intervenção pedagógica.

Nesse sentido, ações específicas podem ser delineadas para cada perfil identificado. Para os Instrumentais, por exemplo, podem ser adotadas medidas que reduzam comportamentos oportunistas, como o aumento da complexidade dos questionários avaliativos, a exigência de acesso mínimo a determinados recursos antes da liberação das avaliações ou a adoção de avaliações que demandem maior articulação conceitual. Para os Engajados e Estratégicos, estratégias de valorização da participação podem ser implementadas, como a oferta de certificados diferenciados que destaquem o desempenho acadêmico, ou incentivos adicionais, a exemplo de benefícios em cursos pagos da própria plataforma ou da instituição ofertante. Já para os Inativos, mecanismos de acompanhamento e intervenção precoce, como notificações personalizadas, mensagens de incentivo ou oferta de apoio pedagógico, podem contribuir para a redução das taxas de evasão.

Por fim, este estudo reforça a relevância do mapeamento de perfis comportamentais como uma estratégia promissora para compreender a complexidade do uso de MOOC e subsidiar a personalização de ambientes de aprendizagem em larga escala. Ao evidenciar que diferentes objetivos coexistem em um mesmo curso, a pesquisa contribui para o avanço teórico e metodológico da área, ao mesmo tempo em que aponta caminhos concretos para o aprimoramento da qualidade e da efetividade dos MOOC.

8. REFERÊNCIAS

- CAGILTAY, N. E.; TOKER, S.; CAGILTAY, K. Exploring MOOC learners' behavioural patterns considering age, gender and number of course enrolments: insights for improving educational opportunities. **Open Praxis**, Barcelona, v. 16, n. 1, p. 1–15, 2024. DOI: <https://10.55982/openpraxis.16.1.579>.
- CHEN, Y.; ZHANG, Y.; LIU, X.; WANG, S. Research on learner emotion–behavior–ability characteristics based on MOOC online education user profiles. **Information Processing & Management**, Amsterdam, v. 62, n. 1, p. 104026, 2025. DOI: <https://10.1016/j.ipm.2024.104026>.
- EL AYADY, A.; DEVANNE, M.; FORESTIER, G.; EL MAWAS, N. Failure risk prediction in a MOOC: a multivariate time series analysis approach. **arXiv**, 2025. Disponível em: <https://arxiv.org/abs/2507.21118>. Acesso em: 30 jan. 2026.
- FELDER, Richard; SILVERMAN, Linda. Learning and Teaching Styles in Engineering Education. **Engineering Education**, [S. l.], v. 78, n. 7, p. 674–681, 1988.
- FERREIRA-SATLER, Mateus *et al.* Fuzzy ontologies-based user profiles applied to enhance e-learning activities. **Soft Computing**, [S. l.], v. 16, n. 7, p. 1129–1141, 2012. Disponível em: <https://doi.org/10.1007/s00500-011-0788-y>.
- FERREIRA, Hiran Nonato M. *et al.* An automatic and dynamic student modeling approach for adaptive and intelligent educational systems using ontologies and Bayesian networks. *In*: 2017, **Proceedings - 2016 IEEE 28th International Conference on Tools with Artificial Intelligence, ICTAI 2016**. [S. l.: s. n.] p. 738–745. Disponível em: <https://doi.org/10.1109/ICTAI.2016.0113>.
- FINNEY, Sara J. *et al.* Exploring profiles of academic help seeking: A mixture modeling approach. **Learning and Individual Differences**, [S. l.], v. 61, n. February 2017, p. 158–171, 2018. Disponível em: <https://doi.org/10.1016/j.lindif.2017.11.011>.
- HAMIM, Touria; BENABBOU, Faouzia; SAEL, Nawal. Student profile modeling: An overview model. *In*: 2019, **ACM International Conference Proceeding Series**. [S. l.: s. n.] p. 1–9. Disponível em: <https://doi.org/10.1145/3368756.3369075>.
- HARTIGAN, J. A.; WONG, M. A. (1979). Algorithm AS 136: A *K-means* clustering algorithm. **Applied Statistics**, v. 28, p. 100–108. Disponível em: <https://doi.org/10.2307/2346830>.
- HUANG, Hao; JEW, Lihjen; QI, Dandan. Take a MOOC and then drop: a systematic review of MOOC engagement pattern and dropout factors. *Heliyon*, Amsterdam, v. 9, n. 4, e15220, 2023. DOI: <https://10.1016/j.heliyon.2023.e15220>.
- KOLB, D. .. **Experiential Learning: Experience as the Source of Learning and Development**. New Jersey: Englewood Cliffs, 1984. *E-book*.
- KORCHI, Adil; ABDELLAH, Mohamed Ben; FEZ, Morocco. A modeling learner approach in a computing environment for human learning based on ontology. **International Journal of**

Computer Engineering & Technology International Journal of Computer Engineering and Technology, [S. l.], v. 6, n. 69, p. 21–31, 2015.

LUNA, J. M.; FARDOUN, H. M.; PADILLO, F.; ROMERO, C.; VENTURA, S. Subgroup discovery in MOOCs: a big data application for describing different types of learners. **arXiv**, 2024. Disponível em: <https://arxiv.org/abs/2403.05555>. Acesso em: 30 jan. 2026.

MARECA, Pilar; BORDEL, Borja. Students Profiles and their Behavior in MOOC Platforms, MIRIADAX Platform. *In*: 2019, Coimbra, Portugal. **14th Iberian Conference on Information Systems and Technologies (CISTI)**. Coimbra, Portugal: [s. n.], 2019. p. 1–6. Disponível em: <https://doi.org/10.23919/CISTI.2019.8760696>.

MUMFORD, A. Honey P. **The Manual of Learning Styles**. [S. l.]: Maidenhead, 1982. *E-book*.

REZGUI, Kalthoum; MHIRI, Hédia; GHÉDIRA, Khaled. Ontology-based e-Portfolio modeling for supporting lifelong competency assessment and development. **Procedia Computer Science**, [S. l.], v. 112, p. 397–406, 2017. Disponível em: <https://doi.org/10.1016/j.procs.2017.08.041>.

RODRIGUES, Rodrigo Lins *et al.* Discovery engagement patterns MOOCs through cluster analysis. **IEEE Latin America Transactions**, [S. l.], v. 14, n. 9, p. 4129–4135, 2016. Disponível em: <https://doi.org/10.1109/TLA.2016.7785943>.

SALLES, I.; BOUHADI, S.; FORESTIER, G.; MOUSSAOUI, A. Interpret3C: interpretable student clustering through individualized feature selection. **arXiv**, 2024. Disponível em: <https://arxiv.org/abs/2407.11979>. Acesso em: 30 jan. 2026.

SHAH, Dhawal. **By The Numbers: MOOCs in 2021. Class Central, 1 dez. 2021**. Disponível em: <https://www.classcentral.com/report/mooc-stats-2021/>. Acesso em: 28 fev. 2026.

SHI, Yuling; PENG, Zhiyong; WANG, Hongning. Modeling student learning styles in MOOCs. *In*: 2017, Singapura. **International Conference on Information and Knowledge Management**. Singapura: [s. n.], 2017. p. 979–988. Disponível em: <https://doi.org/10.1145/3132847.3132965>.

SILVA, J. M. C. da; SANTOS, E. F.; SOUZA, R. M.; OLIVEIRA, D. S. Learner engagement and demographic influences in Brazilian massive open online courses: a case study of the Aprenda Mais platform. **Analytix**, Basel, v. 3, n. 2, p. 178–193, 2024. DOI: <https://10.3390/analytix3020010>.

SNYMAN, Marici; VAN DEN BERG, Geesje. The Significance of the Learner Profile in Recognition of Prior Learning. **Adult Education Quarterly**, [S. l.], v. 68, n. 1, p. 1–16, 2018. Disponível em: <https://doi.org/10.1177/0741713617731809>.

SPRINGER. Strategic and comprehensive: modeling MOOC learner behavior through process mining of learning sequences. *In*: EUROPEAN MOOCS STAKEHOLDERS SUMMIT – EMOOCS 2025. **Cham: Springer**, 2025. p. 65–79.

TAN, Yueying *et al.* Learning Profiles, Behaviors and Outcomes: Investigating International Students' Learning Experience in an English MOOC. *In*: 2018, **Proceedings - 2018 International Symposium on Educational Technology, ISET 2018**.: IEEE, 2018. p. 214–218. Disponível em: <https://doi.org/10.1109/ISET.2018.00055>.

WANG, Rui; CAO, Jie; XU, Yachen; LI, Yanyan. Learning engagement in massive open online courses: a systematic review. **Frontiers in Education**, Lausanne, v. 7, p. 1074435, 2022. DOI: <https://doi.org/10.3389/educ.2022.1074435>.

Submissão: 03/02/2026

Aceito: 02/03/2026