

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Gustavo Alves Kuabara

**Avaliação de Sistema de Recomendação para
Atividades Pedagógicas no Moodle**

Uberlândia, Brasil

2026

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Gustavo Alves Kuabara

**Avaliação de Sistema de Recomendação para Atividades
Pedagógicas no Moodle**

Trabalho de conclusão de curso apresentado
à Faculdade de Computação da Universidade
Federal de Uberlândia, como parte dos requi-
sitos exigidos para a obtenção título de Ba-
charel em Ciência da Computação.

Orientador: Márcia Aparecida Fernandes

Coorientador: Gustavo Prado Oliveira

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2026

Gustavo Alves Kuabara

Avaliação de Sistema de Recomendação para Atividades Pedagógicas no Moodle

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, como parte dos requisitos exigidos para a obtenção título de Bacharel em Ciência da Computação.

Trabalho aprovado. Uberlândia, Brasil, 11 de agosto de 2026:

Márcia Aparecida Fernandes
Orientador

Maria Adriana Vidigal de Lima
Convidado

Fabiano Azevedo Dorça
Convidado

Uberlândia, Brasil
2026

Dedico este trabalho à minha família por estar sempre presente em diversos momentos da minha vida, por me apoiarem no meu aprendizado, na vida profissional, incentivando meus estudos e me permitirem esta oportunidade.

Agradecimentos

Agradeço à minha Orientadora Márcia Aparecida Fernandes e ao meu Coorientador Gustavo Prado Oliveira que me ajudaram e proporcionaram a realização deste trabalho, tirando minhas dúvidas e me guiando para melhores resultados. Ademais agradeço a amigos e colegas que me acompanharam e me apoiaram nessa trajetória, seja dentro da sala de aula ou na vida pessoal.

Declaração de Uso de Inteligência Artificial

Em conformidade com o Código de Conduta para Autores em Publicações da Sociedade Brasileira de Computação (SBC), declara-se o uso de ferramentas de inteligência artificial generativa na preparação deste trabalho. A ferramenta Gemini foi empregada como apoio em: (i) revisão linguística e estilística do texto (clareza, naturalidade, padronização de terminologia e pontuação); e (ii) geração de scripts auxiliares para a produção das figuras a partir dos dados experimentais.

A concepção do estudo, o desenho experimental, a execução da bateria de ensaios, a coleta e a análise dos dados, a interpretação dos resultados e as conclusões são de autoria exclusiva do autor. Nenhum conteúdo gerado por inteligência artificial foi utilizado como fonte primária; o autor revisou integralmente o texto e assume total responsabilidade por seu conteúdo.

Resumo

O trabalho tem como objetivo avaliar a eficácia de um Sistema de Recomendação Educacional (SRE) junto ao Ambiente Virtual de Aprendizagem Moodle, suportado pelo modelo preditivo *Dynamic Key-Value Memory Networks* (DKVMN). Este projeto foca-se na personalização de trajetórias de aprendizagem para o desenvolvimento de competências complexas nos estudantes. A metodologia adotada recorre à técnica de aprendizagem de máquina não supervisionada, utilizando o algoritmo K-means, para segmentar os estudantes em três perfis cognitivos distintos (Baixa Performance, Intermediário e Alta Performance). A avaliação empírica do sistema foi conduzida ao longo de diferentes módulos de aprendizagem, recorrendo a métricas educacionais específicas como o Ganho Normalizado Médio (*Average Normalized Gain* - ANG) e a taxas comportamentais de Ascensão, Resgate e Queda. Os resultados obtidos evidenciam que uma elevada adesão (aderência) às atividades sugeridas promove ganhos pedagógicos expressivos. Destaca-se a eficácia do sistema no resgate de alunos de baixa performance e, no caso dos alunos de excelência, a manutenção do engajamento e prevenção de quedas de rendimento. Conclui-se que a recomendação personalizada impulsionada pelo sistema atua de forma positiva no desenvolvimento de competências, sugerindo-se, para trabalhos futuros, a expansão longitudinal dos módulos de avaliação e a recalibração de métricas iniciais.

Palavras-chave: avaliação, métricas, aprendizado, clusterização, recomendação.

Lista de ilustrações

Figura 1 – Fluxo Geral do Processo de Avaliação.	24
Figura 2 – Gráficos por Módulo com <i>Silhouette Score</i> e suas variações (Inércia) em função de k	40
Figura 3 – Gráfico de Performance x Módulo em Função do Número de Alunos. .	41

Lista de tabelas

Tabela 1 – Métricas	21
Tabela 2 – Comparação sobre o uso de aprendizado de máquina para avaliação de sistemas de recomendação	23
Tabela 3 – Mapeamento das atividades do Moodle e competências associadas. . .	25
Tabela 4 – Tabela de Performance x Módulo em Função do Número de Alunos. . .	29
Tabela 5 – Tabela da Média das Competências em Relação aos Perfis por Módulo.	29
Tabela 6 – Tabela das Taxas por Aderência do Módulo 1 para o Módulo 2.	31
Tabela 7 – Tabela das Taxas por Aderência do Módulo 2 para o Módulo 3.	32
Tabela 8 – Tabela de Ganho Normalizado Médio (ANG) por Aderência do Módulo 1 para o Módulo 2.	32
Tabela 9 – Tabela de Ganho Normalizado Médio (ANG) por Aderência do Módulo 2 para o Módulo 3.	32
Tabela 10 – Tabela de Correlações.	34
Tabela 11 – Resultados das Correlações do Módulo 1 para o Módulo 2.	35
Tabela 12 – Resultados das Correlações do Módulo 2 para o Módulo 3.	36

Lista de abreviaturas e siglas

AVA	Ambiente Virtual de Aprendizado
SRE	Sistema de Recomendação Educacional
DKVMN	<i>Dynamic Key-Value Memory Networks</i>
KT	<i>Knowledge Tracing</i>
RMSE	<i>Root Mean Square Error</i>
AUC	<i>Area Under the ROC Curve</i>
ANG	<i>Average Normalized Gain</i>
ACC	<i>Accuracy</i>
TA	Taxa de Ascensão
TR	Taxa de Resgate
TQ	Taxa de Queda
EDM	Mineração de Dados Educacionais
IFTM	Instituto Federal do Triângulo Mineiro

Sumário

1	INTRODUÇÃO	12
1.1	Estrutura do Trabalho	13
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	Sistemas de Recomendação Educacional (SRE)	15
2.2	Knowledge Tracing e o Modelo DKVMN	16
2.3	Técnicas de Agrupamento de Dados (Clustering)	16
2.4	Métricas de Avaliação em Contextos Educacionais	17
2.5	Mineração de Dados Educacionais (EDM)	17
3	TRABALHOS RELACIONADOS	19
3.1	Fundamentos de Modelagem Cognitiva e Moodle	19
3.2	Métricas para Avaliação de Sistema de Recomendação	20
3.3	Agrupamento e Knowledge Tracing	22
4	DESENVOLVIMENTO	24
4.1	Contextualização do Sistema de Recomendação Avaliado	24
4.2	Clusterização de Estudantes por Competências e Análise de Perfis	26
4.2.1	Preparação de Dados	27
4.2.2	Determinação do Número de Clusters e Avaliação Estatística	27
4.2.3	Clusterização e Resultados	29
4.3	Avaliação do Sistema de Recomendação	30
4.3.1	Avaliação por Utilização de Métricas entre <i>Clusters</i> e <i>Average Normalized Gain</i> (ANG) Sobre a Aderência	30
4.3.2	Avaliação das Correlações das Atividades e Competências entre <i>Clusters</i>	33
4.3.3	Estudo de Casos: Análise Individualizada de Aderência e Correlações	37
4.4	Sugestões para Adaptação do Sistema de Recomendação Dependente dos Resultados da Avaliação	38
5	CONCLUSÃO	42
	REFERÊNCIAS	44

APÊNDICES	46
APÊNDICE A – EXPLORAÇÃO DE ALGORITMOS BASEADOS EM DENSIDADE	47

1 Introdução

Diante do Ambiente Virtual de Aprendizado (AVA) Moodle, é possível observar as diversas atividades que ele proporciona e sua ampla utilização para os professores e estudantes. Entretanto, seria de interesse o uso do Moodle não apenas como um AVA para postagens de atividades avaliativas, mas como um ambiente para o desenvolvimento das diferentes competências que um aluno possui tendo em vista a quantidade de funções e diferentes feedbacks que o ambiente proporciona. Dito isso, o aluno de doutorado da Universidade Federal de Uberlândia Gustavo Prado Oliveira desenvolveu um sistema de recomendação de atividades do Moodle a partir da ligação das características do sujeito complexo (CSC) termo usado por Edgar Morin ([MORIN, 2000](#)) definindo as mesmas como uma série de competências e com ações pedagógicas ([OLIVEIRA; FERNANDES; MAISSIAT, 2022](#)) e então relacionando-as através de pesquisas feitas com educadores às atividades do Moodle ([OLIVEIRA; FERNANDES; MAISSIAT, 2023](#)). Nesse sentido, o presente estudo busca verificar se as recomendações propostas alinham-se ao perfil dos discentes, assegurando a efetividade da personalização de atividades no ambiente Moodle.

Com a avaliação das recomendações será possível distinguir as atividades relacionadas a certas competências, e então utilizar as recomendações de atividades pedagógicas no Moodle para desenvolvê-las, com a possibilidade de melhorar aspectos do pensamento crítico do aluno e notas. Tais resultados podem ser observados no documento de qualificação de Gustavo Prado Oliveira ([OLIVEIRA, 2026](#)), em uma das instâncias de coleta de dados feitas, em um total de 42 estudantes, 26 obtiveram melhora em pelo menos uma das competências relacionadas as atividades recomendadas. Assim, o tema traz uma relevância na educação para educadores e alunos nos quais interagem com o AVA do Moodle e similares.

Dada a relevância do tema, é possível observar trabalhos relacionados ao mesmo, destaca-se mais uma vez os artigos ([OLIVEIRA; FERNANDES; MAISSIAT, 2022](#)) ([OLIVEIRA; FERNANDES; MAISSIAT, 2023](#)) e qualificação ([OLIVEIRA, 2026](#)) que são a base inicial do projeto, os quais retratam desde a nomeação de competências para as atividades do Moodle, coleta de dados acerca das notas das atividades em três módulos ministrados em disciplinas e demonstração dos resultados. Além disso, ressaltam-se os artigos ([CABINESS; DONOVAN; GREEN, 2013](#)), tratando-se de um projeto feito com estudantes utilizando-se de uma wiki colaborativa entre eles para o desenvolvimento do pensamento crítico, habilidades e reflexão, tendo similar a atividade de wiki proposta pelo Moodle e por [Apino e Retnawati \(2017\)](#), o qual se baseia no uso de testes matemáticos para o desenvolvimento de *higher order thinking skills* (HOTS), mostrando a importância de atividades para o pensamento crítico e encorajamento para os estudos.

Objetivo Geral: Avaliar um sistema de recomendação educacional para a plataforma Moodle.

Objetivos Específicos: Para o alcance do objetivo geral, estabelecem-se os seguintes objetivos específicos:

1. Estudar como avaliar sistemas educacionais: formas, métodos, métricas
2. Definir as métricas para avaliar o sistema proposto
3. Estudar algoritmos de aprendizagem de máquina que possam ser aplicados neste contexto
4. Coletar de dados e adequação dos mesmos ao algoritmo escolhido
5. Realizar experimentos e avaliar o sistema utilizado

Sendo assim, para alcançar tais objetivos, este projeto caracteriza-se como um trabalho científico, utilizando-se de técnicas de aprendizagem de máquina, métodos avaliativos e estatística embasados em literatura consolidada.

1.1 Estrutura do Trabalho

Para melhor compreensão do leitor, este trabalho está organizado em quatro capítulos principais, estruturados da seguinte forma:

- **Capítulo 1 – Introdução:** Contextualiza o uso de Ambientes Virtuais de Aprendizagem (AVAs), em especial o Moodle, apresenta a motivação pedagógica fundamentada nas competências do Sujeito Complexo, a justificativa e os objetivos geral e específicos da pesquisa.
- **Capítulo 2 – Fundamentação Teórica:** Aborda os pilares teóricos do trabalho, incluindo Sistemas de Recomendação Educacional (SRE), algoritmos de traçagem de conhecimento (*Knowledge Tracing*) focando no modelo *Dynamic Key-Value Memory Networks* (DKVMN), técnicas de agrupamento de dados (*K-means*) e métricas de avaliação em contextos educacionais, no caso, o *Average Normalized Gain* (ANG).
- **Capítulo 3 – Trabalhos Relacionados:** Detalha trabalhos relacionados à presente avaliação, como o trabalho base alvo da avaliação de [Oliveira \(2026\)](#). Além dos trabalhos de [Akhadam et al. \(2025\)](#) e [Hake \(1998\)](#) que destacam as métricas de avaliação, bem como os trabalhos de [Zhang et al. \(2017\)](#) e [Jain \(2010\)](#) que dizem respeito ao modelo de *knowledge tracing* e agrupamento, respectivamente.

- **Capítulo 4 – Desenvolvimento:** Detalha a arquitetura da solução, apresentando o fluxo geral do processo de avaliação, a etapa de Mineração de Dados Educacionais (EDM), a matriz de correlação de competências pedagógicas e os procedimentos de separação e avaliação dos grupos. Além disso, apresenta e discute os achados empíricos (evolução de *clusters*, taxas de resgate e ascensão por aderência) e ganhos cognitivos e efeitos indiretos por atividade. Por fim, se aprofunda nos resultados exaltando estudo de casos e realizando sugestões ao sistema de recomendação.
- **Capítulo 5 – Conclusão:** Encerra com a síntese do trabalho e dos resultados, juntamente às conclusões tiradas e sugestões de trabalhos futuros.

2 Fundamentação Teórica

Neste capítulo, são abordados os conceitos fundamentais que sustentam o desenvolvimento e a avaliação deste projeto. Inicialmente, discute-se o paradigma dos Sistemas de Recomendação Educacional (SRE), diferenciando-os dos sistemas comerciais. Em seguida, introduz-se o conceito de Traçagem de Conhecimento (Knowledge Tracing) e o modelo Dynamic Key-Value Memory Networks (DKVMN), métodos cruciais para a modelação da aprendizagem. Adicionalmente, exploram-se as técnicas de agrupamento de dados (Clustering) para a identificação de perfis de estudantes e, por fim, definem-se as métricas de avaliação adequadas a contextos educacionais.

2.1 Sistemas de Recomendação Educacional (SRE)

Sistemas de Recomendação (SR) são ferramentas amplamente utilizadas para auxiliar os usuários a encontrar itens relevantes em um vasto volume de informações. Em contextos comerciais, como plataformas de *e-commerce* ou de *streaming* (e.g., Netflix, Spotify), o objetivo principal destas ferramentas é maximizar o consumo e o engajamento do usuário, sugerindo produtos com base em preferências passadas. No entanto, a aplicação destas tecnologias no domínio da educação requer uma mudança substancial de paradigma.

Os Sistemas de Recomendação Educacional (SRE) diferenciam-se essencialmente pelo seu propósito pedagógico. Conforme analisado no estudo de revisão conduzido por [Klašnja-Milićević, Ivanović e Nanopoulos \(2015\)](#), enquanto os sistemas comerciais visam a conversão de vendas, os SREs têm como foco central a personalização da aprendizagem. Os autores destacam que os estudantes em ambientes *e-learning* possuem diferenças significativas em termos de conhecimento prévio, objetivos, capacidades cognitivas e estilos de aprendizagem. Dessa forma, o papel do SRE é atuar como um facilitador adaptativo.

Em vez de produtos de consumo, estes sistemas recomendam “Objetos de Aprendizagem” — tais como atividades pedagógicas, leituras ou exercícios em um Ambiente Virtual de Aprendizagem (AVA) como o Moodle — com o intuito de suprir lacunas de conhecimento e desenvolver competências específicas de forma individualizada ([KLAŠNJA-MILIĆEVIĆ; IVANOVIĆ; NANOPOULOS, 2015](#)). A eficácia dessa abordagem tem se mostrado crescente na literatura. O estudo de mapeamento sistemático conduzido por [Rivera, Tapia-Leon e Lujan-Mora \(2018\)](#), por exemplo, mostra a integração de SREs em AVAs nos últimos anos, evidenciando o impacto positivo na criação de percursos de estudo que respeitam o ritmo e as necessidades cognitivas de cada aluno.

2.2 Knowledge Tracing e o Modelo DKVMN

Para que um Sistema de Recomendação Educacional seja verdadeiramente eficaz, é necessário compreender o estado atual do conhecimento do estudante. É neste contexto que surge a Traçagem de Conhecimento ou *Knowledge Tracing* (KT), uma área que se dedica a modelar e rastrear a evolução do domínio que um aluno tem sobre determinados conceitos ao longo do tempo, baseando-se no seu histórico de interações e no desempenho em atividades passadas.

Uma das abordagens modernas e altamente eficazes para este fim é o modelo DKVMN (Dynamic Key-Value Memory Networks), introduzido por [Zhang et al. \(2017\)](#). O DKVMN é um modelo clássico de Deep Knowledge Tracing que se destaca por utilizar uma arquitetura de memória externa dividida em duas matrizes: uma matriz estática *Key* que armazena os conceitos de conhecimento, e uma matriz dinâmica *Value* que atualiza e armazena o nível de domínio do estudante sobre esses mesmos conceitos à medida que ele resolve exercícios.

Esta separação permite ao modelo não só prever com maior precisão se o aluno irá acertar ou errar um exercício futuro, mas também interpretar quais os conceitos subjacentes que o aluno domina ou nos quais apresenta dificuldades, fornecendo dados valiosos para a personalização das recomendações no Moodle.

2.3 Técnicas de Agrupamento de Dados (Clustering)

A compreensão do comportamento dos estudantes e do impacto das atividades educacionais pode ser aprofundada através da utilização de algoritmos de Aprendizagem de Máquina Não Supervisionada, especificamente as técnicas de agrupamento (*Clustering*). De acordo com [Jain \(2010\)](#), o agrupamento de dados é o processo de organização de objetos em grupos cujos membros são semelhantes entre si, diferenciando-se consideravelmente dos membros de outros grupos.

No contexto deste trabalho, o *clustering* permite agrupar os alunos com base no seu desempenho em diferentes competências e atividades. Embora existam diversas abordagens descritas na literatura (como métodos hierárquicos e baseados em densidade), o algoritmo K-means destaca-se como uma das soluções mais consolidadas e eficientes para a partição de dados em k grupos distintos. A aplicação destas técnicas facilita a visualização e a inferência de padrões de aprendizagem, permitindo, por exemplo, correlacionar clusters de alunos com necessidades específicas da recomendação.

2.4 Métricas de Avaliação em Contextos Educacionais

A avaliação de um modelo preditivo ou de um sistema de recomendação no contexto educacional exige a seleção de métricas que reflitam adequadamente o seu propósito. Embora métricas focadas no erro probabilístico, como o RMSE (*Root Mean Square Error*) (HERLOCKER et al., 2004), sejam tradicionais na avaliação de sistemas de recomendação comerciais, elas raramente constituem a métrica principal para modelos de avaliação de aprendizagem como o DKVMN.

Na literatura especializada de traçagem de conhecimento, os modelos são avaliados principalmente através de métricas de classificação, conforme estabelecido pelo trabalho de Zhang et al. (2017):

Area Under the ROC Curve (AUC): Considerada a métrica de referência para Knowledge Tracing, o AUC avalia a capacidade do modelo em distinguir corretamente entre um aluno que domina o conceito (probabilidade de acerto) e um que não domina (probabilidade de erro).

Accuracy (ACC): Mede a proporção geral de predições corretas (acertos ou erros) realizadas pelo modelo em relação ao total de interações.

Por fim, no que concerne à avaliação quantitativa da evolução pedagógica, destaca-se o uso de métricas como o *Average Normalized Gain* (ANG), estabelecido por Hake (1998), que permite quantificar o impacto de uma metodologia (ou, neste caso, de uma recomendação) no ganho efetivo de aprendizagem do estudante.

2.5 Mineração de Dados Educacionais (EDM)

A Mineração de Dados Educacionais (*Educational Data Mining* – EDM) consiste na aplicação de técnicas de aprendizado de máquina, estatística e mineração de dados sobre registros gerados em ambientes de aprendizagem (ROMERO; VENTURA, 2010). De acordo com Romero e Ventura (2010), o ciclo de EDM envolve fases iterativas de extração de dados, pré-processamento, aplicação de algoritmos de descoberta de padrões (como o agrupamento) e a subsequente interpretação pedagógica dos resultados.

Ademais, destaca-se o trabalho de Dutt, Ismail e Herawan (2017), que aponta as técnicas de agrupamento (*clustering*) e de classificação entre as mais utilizadas na Mineração de Dados Educacionais para a análise de perfis estudantis, além de evidenciar o Moodle como uma das principais fontes de extração desses dados.

Por fim, aprofundando o aspecto metodológico, Aldowah, Al-Samarraie e Fauzy (2019) utilizam a mineração de dados em conjunto com técnicas de agrupamento para aprimorar a análise de perfis de estudantes. Segundo os autores, essa abordagem per-

mite estruturar e analisar diversos padrões comportamentais, desde a predição de possível evasão na disciplina até a identificação de alunos com alto potencial de crescimento acadêmico.

No presente trabalho, o arcabouço metodológico de EDM é utilizado no Desenvolvimento para transformar os resultados brutos do sistema de recomendação em características interpretáveis a partir do agrupamento dos discentes por perfis de desempenho e a análise da evolução cognitiva.

3 Trabalhos Relacionados

Com o objetivo de avaliar um Sistema de Recomendação Educacional (SRE), a busca por trabalhos relacionados torna-se fundamental a fim de interligar aspectos já presentes na literatura e demonstrar as pesquisas que atuam como base para o presente estudo.

Dessa forma, serão abordados estudos relacionados à Modelagem Cognitiva e Moodle, Métricas para Avaliação de Sistemas de Recomendação, Agrupamento e *Knowledge Tracing*, evidenciando a contribuição desta pesquisa para o desenvolvimento acadêmico da área.

3.1 Fundamentos de Modelagem Cognitiva e Moodle

O principal ponto deste trabalho reside na pesquisa de doutoramento apresentada na qualificação por [Oliveira \(2026\)](#), que estabeleceu um protocolo formal e validado para a recomendação personalizada de atividades em Educação a Distância. O trabalho, que utilizou como base a Teoria da Complexidade de Edgar Morin ([MORIN, 2000](#)) para definir o “sujeito complexo” como um conjunto de competências dinâmicas, operou uma ponte entre a pedagogia e a inteligência artificial.

Metodologicamente, o autor utilizou-se do conceito expresso em Taxonomia Digital de Bloom ([CHURCHES, 2008](#)) para rotular atividades do Moodle, criando um mapeamento técnico-pedagógico onde características de ferramentas (como Fóruns, Wikis e Lições) foram correlacionadas a processos cognitivos de alta ordem, conforme a caracterização proposta por [Pacheco e Herrera \(2021\)](#). O processo de extração de dados envolveu a coleta sistemática de logs do Moodle, onde algoritmos de pré-processamento foram aplicados para categorizar as interações dos usuários em termos de proficiência nas competências mapeadas.

A validação da proposta foi conduzida através de um experimento com 42 estudantes, no qual o modelo de recomendação — treinado sobre os dados rotulados — demonstrou eficácia quantitativa: 26 estudantes apresentaram melhorias mensuráveis em competências específicas. Tal resultado não apenas sugere a validação do método de rotulação, mas provê a matriz de dados (notas por competência) que serve como insumo para os modelos de *Knowledge Tracing* explorados no presente projeto. Portanto, a pesquisa de [Oliveira \(2026\)](#) não é apenas uma referência teórica, mas o trabalho central para a aquisição de dados e definição de competências que viabiliza a análise aqui proposta.

3.2 Métricas para Avaliação de Sistema de Recomendação

Introduzindo diversos pontos de sistemas de recomendação o artigo ([AKHADAM et al., 2025](#)) se demonstra completo, em que, revisa sobre os principais algoritmos, métricas para avaliação de sistema de recomendação, *frameworks*, avaliação de ferramentas, datasets e propõem a implementação de diversos algoritmos juntamente às ferramentas, datasets, realizando testes, avaliando e fazendo análise dos resultados, com o objetivo de determinar a eficiência dos algoritmos e compará-los.

No artigo, as seguintes ferramentas foram utilizadas: Microsoft Recommenders, Turi Create e Cornac. Em seguida, são apresentadas as métricas para avaliação dos algoritmos a serem utilizadas, sendo encontradas na Tabela 1.

Ademais, os datasets utilizados foram: Movie Lens: do GroupLens Research, notas de filmes de 0,5 a 5, Jester: do Jester Research project da Universidade da California, notas de 100 piadas de -10.00 a 10.00, Bookcrossing: coletado por Cai-Nicolas Ziegler do Book-Crossing community, notas de 1-10 explícitas ou 0-1 implícitas, Epinions: coletado por Paulo Massa, do Epinions.com, notas de 1-5 de produtos. O Pré-Processamento foi feito em todas as bases de dados visto irregularidades e dados não concretos. Por fim dentre as técnicas utilizadas podemos citar: *Collaborative Filtering (User-User, Item-Item, Alternating Least Squares, Random Forest Regression)*, *Matrix Factorization (Singular Value Decomposition, Bayesian Personalized Ranking)*, *Smart Adaptative Recommendations, Item Similarity, K-Nearest-Neighbors (User-Based, Item-Based)*. Com base nas ferramentas e algoritmos citados, foram feitos diversos testes a partir das bases de dados para cada algoritmo e então avaliado através das métricas.

Como resultados, cada ferramenta apresenta pequenos benefícios em relação as outras, em geral, nenhuma se destacando excessivamente. No entanto, algoritmos baseados em fatoração de matrizes, como *Singular Value Decomposition*, demonstraram desempenho robusto e versátil, tornando-os escolhas confiáveis para diversos cenários de recomendação. Métodos baseados em vizinhança, como UserKNN e ItemKNN, apresentaram resultados mistos, destacando-se em ferramentas como Lenskit, mas apresentando desempenho inferior em outras, como Cornac, demonstrando uma possível diferença de performance sobre diferentes ferramentas.

Tal artigo pode relacionar-se ao projeto por apresentar de maneira geral e empírica testes e avaliações feitas em sistemas de recomendação, bem como diversas métricas que possivelmente podem ser utilizadas para avaliar um sistema de recomendação, bem como uma análise dos resultados bem elaborada, trazendo aspectos a serem analisados ao apresentar de resultados do projeto.

Além disso, como base para métricas educacionais relacionados à aprendizagem, pode-se citar o artigo ([HAKE, 1998](#)) que possui como objetivo verificar a eficiência dos

Tabela 1 – Métricas

Symbol	Metric	Formula
MAE	Mean Absolute Error	$\frac{1}{N} \sum_{i=1}^N r_{ui} - \tilde{r}_{ui} $
RMSE	Root Mean Squared Error	$\frac{1}{N} \sum_{i=1}^N \sqrt{(r_{ui} - \tilde{r}_{ui})^2}$
PC	Pearson Correlation	$\frac{\sum_{i=1}^N (r_{ui} - \bar{r}_u)(\tilde{r}_{ui} - \bar{\tilde{r}}_u)}{\sqrt{\sum_{i=1}^N (r_{ui} - \bar{r}_u)^2 (\tilde{r}_{ui} - \bar{\tilde{r}}_u)^2}}$
SC	Spearman Correlation	$\frac{\sum_{i=1}^N (r_i - \bar{r}_u)(\hat{r}_i - \bar{\hat{r}}_u)}{\sqrt{\sum_{i=1}^N (r_i - \bar{r}_u)^2 (\hat{r}_i - \bar{\hat{r}}_u)^2}}$
P	Precision	$\frac{ C_u \cap R_u }{ R_u }$
R	Recall	$\frac{ C_u \cap R_u }{ C_u }$
F1	F1 score	$\frac{2PR}{P+R}$
NDCG	Normalized Discounted Cumulative Gain	$\frac{\sum_{i=1}^b R_i + \sum_{i=b+1}^N \frac{R_i}{\log b r_i}}{DCG_{max}}$
RBP	Rank-Biased Precision	$(1 - p) \sum_{i=1}^N (R_i \times p^{i-1})$
RR	Recovery Rate	$\frac{1}{ N_R } \sum_{i \in N_R} \frac{r_i}{C_i}$
ID	Intra-list Diversity	$\frac{\frac{1}{N(N-1)} \times \sum_{i \neq j} s(I_i, I_j)}{m}$
H	Hamming distance (diversity)	$\frac{1}{m(m-1)} \sum_{u=1}^m \sum_{u' \neq u} 1 - (\frac{Q_{uu'}}{N})$
Pop	Popularity	$\frac{1}{m \times N} \sum_{u=1}^m \sum_{i=1}^N d_i$
C	Catalogue Coverage	$\frac{l_c}{n}$
EC	Entropy Coverage	$-\sum_{i=1}^m p_i \log_2 p_i$

Fonte: (AKHADAM et al., 2025)

métodos de ensino *interactive-engagement* (ie) em comparação aos métodos tradicionais, e avaliá-los utilizando-se da métrica *Average Normalized Gain* (ANG).

Em sua metodologia o artigo destaca o uso de uma amostra de mais de 6000 estudantes de 62 cursos diferentes de diferentes instituições acadêmicas nos Estados Unidos, em que foi utilizado das ferramentas de teste para avaliação dos estudantes o Force Concept Inventory (FCI) e Mechanics Diagnostic (MD) e se utilizando da métrica ANG para medir o aprendizado dos alunos.

Por fim o artigo apresenta como resultado o *Average Normalized Gain* (atribuindo g para o valor) dos dois métodos de ensino, sendo o g para o ensino tradicional de $0,23 \pm 0,04$ e para o ensino ie de $0,48 \pm 0,14$, demonstrando um maior ganho conceitual.

Tal artigo se relaciona ao projeto atual, embora antigo, demonstrando um dos primeiros usos de métricas de avaliação no contexto educacional, sendo o impulsionador para o uso do ANG, assim, sendo considerado uma base no contexto de avaliação educacional.

3.3 Agrupamento e Knowledge Tracing

Neste projeto, a recomendação a ser avaliada fundamenta-se na integração de modelos de *Knowledge Tracing* e possui como uns dos principais pontos as técnicas de agrupamento, para analisar as recomendações pedagógicas.

O modelo *Dynamic Key-Value Memory Networks* (DKVMN), conforme proposto por [Zhang et al. \(2017\)](#), foi utilizado para realizar a modelagem do perfil cognitivo do estudante. Diferente de abordagens estáticas, o DKVMN mantém uma estrutura de memória externa capaz de atualizar, a cada interação, o nível de proficiência do aluno em competências específicas. Esta capacidade diagnóstica dinâmica permite que o sistema identifique lacunas de aprendizado com base no histórico de desempenho, fornecendo os parâmetros necessários para adaptar o aluno nas atividades.

Paralelamente, a estruturação dos dados pedagógicos é realizada por meio de algoritmos de agrupamento (*clustering*). Conforme [Jain \(2010\)](#), a clusterização é uma ferramenta essencial para a descoberta de estruturas em dados complexos; no escopo deste trabalho, o K-means é empregado para organizar as atividades em grupos com características similares. Esta organização não apenas facilita a visualização de possíveis relações entre atividades e desempenho, mas pode permitir visualizar relação entre competências e a maneira que cada atividade interage sobre as mesmas.

Do ponto de vista matemático e computacional, o algoritmo K-means particiona o conjunto de dados observados em k grupos distintos, com o objetivo de minimizar a variância intra-cluster. Esse processo de otimização é formalizado pela minimização da função objetivo de Soma dos Quadrados Intra-Cluster (do inglês, *Within-Cluster Sum of*

Squares - WCSS), dada pela Equação 3.1:

$$J = \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (3.1)$$

Onde k representa o número predefinido de *clusters*, S_i é o conjunto de dados pertencentes ao i -ésimo *cluster*, x é a observação individual (ponto de dado representativo da métrica do aluno ou da atividade) e μ_i é o centroide (média matemática) do respectivo grupo. Ao aplicar essa formulação matemática, o sistema consegue não apenas facilitar a visualização de possíveis correlações entre atividades e desempenho, mas também identificar de forma quantitativa como cada atividade atua no desenvolvimento de diferentes competências.

Dessa forma, a integração entre a recomendação feita pelo DKVMN e a segmentação das recomendações via *clustering* constitui centro deste projeto. A avaliação da eficácia dessas recomendações deve ser focada em avaliar o impacto pedagógico das sugestões geradas, focando na evolução da trajetória de aprendizagem do estudante.

Tabela 2 – Comparação sobre o uso de aprendizado de máquina para avaliação de sistemas de recomendação

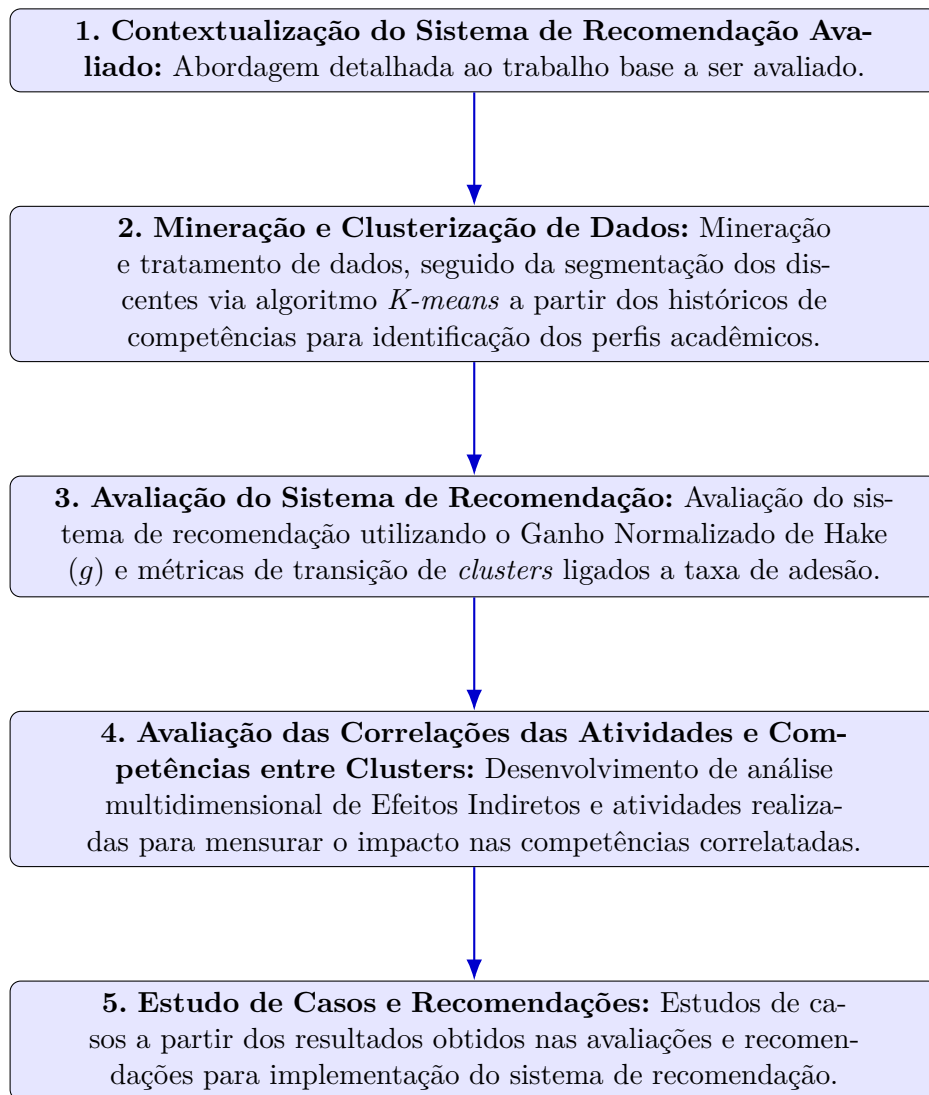
Trabalho 1	Trabalho 2	Trabalho 3
Rotulação de habilidades e atividades	Métricas e avaliação de resultados de sistemas de recomendação e quantitativa educacional	Uso de algoritmos não supervisionados para agrupamento dos dados paralelo ao DKVMN
Base das notas por competências a serem analisadas	Entendimento geral sobre sistemas de recomendação e para avaliação no contexto educacional	Agrupamento para avaliação dos dados

Fonte: Elaborada pelo autor (2026).

4 Desenvolvimento

A partir do conjunto de referências bibliográficas revistas e para mensurar o impacto do sistema de recomendação no contexto pedagógico, estabeleceu-se um fluxo metodológico seguindo em parte o processo de Mineração de Dados Educacionais (EDM), e ilustrado no fluxo metodológico 1 a seguir:

Figura 1 – Fluxo Geral do Processo de Avaliação.



Fonte: Elaborada pelo autor (2026).

4.1 Contextualização do Sistema de Recomendação Avaliado

Antes de detalhar os procedimentos avaliativos, é necessário apresentar a estrutura e o funcionamento do sistema de recomendação alvo deste estudo. O sistema, apresen-

tado no documento de qualificação (OLIVEIRA, 2026), fundamenta-se na premissa de que o Ambiente Virtual de Aprendizagem (AVA) Moodle transcende a função de mero repositório de conteúdos, atuando como um espaço propício para o desenvolvimento de competências complexas nos estudantes.

A lógica de recomendação pedagógica foi estruturada a partir do mapeamento das características do sujeito complexo, embasada na teoria de Morin (2000). Nesse modelo, atividades do Moodle (como Fórum, Wiki, Questionários e Lições) foram catalogadas, rotuladas e associadas a competências cognitivas, tais como Autonomia, Cooperação, Resiliência e Transdisciplinaridade. Esse mapeamento entre as ferramentas da plataforma e as competências é expresso na Tabela 3, que apresenta a rotulação entre competências e atividades de maneira inicial e final.

Tabela 3 – Mapeamento das atividades do Moodle e competências associadas.

Atividades	Glossário		Base de Dados		Chat		Escolha		Fórum		Workshop		Lição		Wiki	
	Inicial	Final	Inicial	Final	Inicial	Final	Inicial	Final	Inicial	Final	Inicial	Final	Inicial	Final	Inicial	Final
Resiliência													X	X		
Autonomia		X		X				X	X	X		X		X		X
Cooperação	X	X	X			X			X	X	X	X				X
Metacognição		X		X			X	X							X	
Transdisciplinaridade						X				X						X
Flexibilidade					X	X			X				X	X		
Criatividade	X		X	X				X			X					
% Equivalência	41,50%		41,50%		66,50%		66,50%		66,60%		50,00%		83,30%		0,00%	

Fonte: Adaptado de (OLIVEIRA, 2026).

A coleta de dados empíricos acompanhou aproximadamente 42 estudantes pertencentes a uma mesma turma do curso de Licenciatura em Computação do Instituto Federal do Triângulo Mineiro (IFTM), ao longo de três semestres acadêmicos consecutivos. Em cada período letivo, a turma esteve matriculada em uma disciplina distinta do programa curricular: Estrutura de Dados, Programação Orientada a Objetos e Noções de Inteligência Artificial.

Cada disciplina foi organizada em três módulos didáticos. Durante e ao final de cada módulo, registraram-se e analisaram-se as interações dos estudantes no AVA Moodle, as recomendações emitidas pelo sistema, o nível de adesão dos discentes às atividades sugeridas e a consequente evolução de suas competências cognitivas.

A dinâmica de aplicação do sistema e a coleta de dados empíricos foram estruturadas ao longo de três módulos instrucionais sequenciais (módulo 1, módulo 2 e módulo 3), funcionando da seguinte forma:

- **Módulo 1 (Linha de Base / Diagnóstico):** Atuou como a etapa inicial de calibração. Neste primeiro momento, registraram-se os desempenhos e interações dos estudantes sobre as atividades para estabelecer o perfil cognitivo inicial de cada aluno.

- **Módulos 2 e 3 (Módulos sobre Recomendações):** A partir do módulo 2, com o histórico do módulo 1 já processado, o sistema passou a atuar com base em recomendações. Com base nas fragilidades detectadas pelo DKVMN, atividades personalizadas foram recomendadas aos alunos para execução nos módulos 2 e 3. Essa divisão em etapas temporais permitiu acompanhar a evolução do desempenho acadêmico ao longo das transições entre os módulos e analisar impacto pedagógico proporcionado pela adesão às recomendações.

Para a efetivação das recomendações foi utilizado do Modelo DKVMN que ao identificar uma lacuna de aprendizado ou um desempenho subótimo em determinada competência, o modelo atua ativamente gerando a recomendação da próxima atividade pedagógica mais adequada para suprir essa deficiência e impulsionar o perfil do estudante nas competências de forma ao ter uma maior variação.

Deve-se lembrar que os dados utilizados para a avaliação foram coletados durante a aplicação do modelo de recomendação desenvolvido na pesquisa de [Oliveira \(2026\)](#), desenvolvendo sobre os dados coletados foram registrados e utilizados:

- Recomendações de atividades para os alunos nos módulos 1 e 2, recomendando atividades seguida das competências a aprimorar e sua pontuação prevista.
- IDs, performance e notas das competências dos alunos nos três módulos apresentados.
- Aderência dos alunos às recomendações, através da porcentagem realizada pelo mesmo em cada módulo.
- Atividades realizadas pelos alunos em cada módulo.

4.2 Clusterização de Estudantes por Competências e Análise de Perfis

O passo fundamental para uma avaliação contextualizada do sistema de recomendação reside na identificação dos diferentes perfis de aprendizagem presentes nos estudantes. Na literatura de Mineração de Dados Educacionais (*Educational Data Mining* - EDM), a aplicação de algoritmos de agrupamento não supervisionado é amplamente utilizada como uma estratégia eficaz para lidar com a heterogeneidade das turmas, permitindo que intervenções pedagógicas e recomendações sejam avaliadas de acordo com as necessidades específicas de cada grupo de alunos ([ROMERO; VENTURA, 2010](#)).

Nesta perspectiva, a clusterização deve permitir mensurar o impacto das recomendações geradas pelo modelo DKVMN ao longo dos módulos de aprendizagem.

4.2.1 Preparação de Dados

Para a caracterização dos perfis, foram extraídos dados referentes ao desempenho dos alunos em atividades práticas e ao somatório de competências (*skills*) mapeadas no ambiente virtual de aprendizagem Moodle.

Considerando que o algoritmo K-means baseia-se no cálculo da distância Euclidiana para determinar a similaridade entre os objetos do conjunto de dados, a escala das variáveis exerce influência direta na formação dos agrupamentos. Atributos com maiores amplitudes numéricas poderiam enviesar o cálculo da distância, descaracterizando o peso de competências (HAN; KAMBER; PEI, 2011).

Para mitigar este efeito e garantir a integridade matemática do modelo, aplicou-se a técnica de padronização *z-score* (ou normalização por escala padrão), descrita pela Equação 4.1:

$$z = \frac{x - \mu}{\sigma} \quad (4.1)$$

Onde x representa o valor original do atributo, μ corresponde à média do atributo no conjunto de dados e σ denota o desvio padrão. Este processo projeta os dados em uma escala comum com média zero e variância unitária, assegurando que todas as competências analisadas contribuam de maneira equitativa para a definição das distâncias entre os estudantes.

Além disso, para não interferir de forma negativa na avaliação, foram retirados os alunos que não tiveram participação alguma ao longo dos três módulos, pois os dados referentes a eles não possuem valor educacional sobre o sistema de recomendação a ser avaliado.

Outro ponto a ser contabilizado no tratamento dos dados seria o carregamento de bagagens dos alunos, ou seja, para os alunos que realizaram as atividades mas diante das recomendações propostas pelo algoritmo não obtiveram nota em certas competências, por conta dessas atividades não terem relação com tais competências, foi repassado os valores das competências adquiridos em módulos anteriores, visto que, o aluno não “desapareceu” com o conhecimento adquirido por ele, logo, foi decidido em conjunto ao coorientador a repassar esses valores.

4.2.2 Determinação do Número de Clusters e Avaliação Estatística

Um dos desafios ao uso do algoritmo K-means é a necessidade de definição prévia do número de agrupamentos (k). A escolha arbitrária deste parâmetro pode resultar em partições subótimas, mascarando nuances pedagógicas relevantes ou fragmentando excessivamente grupos coesos (JAIN, 2010).

Para determinar o valor ideal de k aplicável ao conjunto de dados de desempenho dos estudantes, utilizou-se o coeficiente de silhueta (*Silhouette Score*) como métrica de validação interna. Proposto por [Rousseeuw \(1987\)](#), o coeficiente de silhueta avalia tanto a coesão interna do cluster (distância média entre pontos do mesmo grupo) quanto a sua separabilidade (distância média em relação ao cluster vizinho mais próximo). O índice de silhueta $s(i)$ para um objeto individual i é formalizado pela Equação 4.2:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4.2)$$

Onde $a(i)$ é a distância média de i em relação a todos os outros pontos no mesmo cluster, e $b(i)$ é a menor distância média de i em relação aos pontos de qualquer outro cluster do qual i não seja membro. O valor do coeficiente varia de -1 a $+1$, em que valores próximos a 1 indicam que o ponto está muito bem alocado ao seu cluster e distante dos demais.

Além do *Silhouette Score*, para definir um valor ideal ao k , foi empregado o Método do Cotovelo (*Elbow Method*), amplamente fundamentado na literatura para a otimização de partições no algoritmo K-means ([KODINARIYA; MAKWANA, 2013](#)). Este método heurístico avalia a variação dos resultados obtidos pela métrica de Inércia (Soma dos Quadrados Intra-Cluster – WCSS, detalhada na Equação 3.1) em função do aumento progressivo de k . O ponto de inflexão da curva gerada, assemelhando-se a um cotovelo, indica o número ótimo de clusters, estabelecendo o limite onde o acréscimo de novos agrupamentos não resulta em ganhos significativos de variância explicada.

Dito isso, para obtenção de um k ideal, a implementação do algoritmo k-means se trouxe necessária, tendo como base as notas das competências dos alunos por módulo e então retirados os valores das métricas, e apresentadas no gráfico 2 a seguir:

A partir da análise conjunta dessas duas métricas, determinou-se empiricamente que o valor $k = 3$ configura a partição ideal para o estudo. Esta configuração apresenta um equilíbrio: em tentar otimizar os resultados quantitativos das métricas de validação interna e, simultaneamente, garante uma alta interpretabilidade pedagógica dentro do contexto educacional. Com base nisso, os estudantes foram segmentados nos seguintes perfis:

- **Baixa Performance**
- **Intermediário**
- **Alta Performance**

Cabe esclarecer que o nome dado aos grupos são para uma melhor interpretabilidade pedagógica, por exemplo, nos resultados da clusterização, há alunos com notas

maiores em certas competências e bem menores em outras que acabaram entrando no grupo Intermediário e casos similares foram para o grupo Alta Performance.

4.2.3 Clusterização e Resultados

Encontrado um k ideal a ser trabalhado, e segmentado os estudantes em três perfis, deve-se utilizar do algoritmo k-means para separar os estudantes nesses grupos.

Usando os dados tratados, juntamente a biblioteca sklearn, que possui a função KMeans(), e utilizando como parâmetros: $k = 3$ (número de clusters), $random_state = 42$ e $n_init = 10$ (estabelecidos como padrão na utilização da função) retirou-se os seguintes resultados do algoritmo apresentados na tabela 4 e do gráfico 3.

Tabela 4 – Tabela de Performance x Módulo em Função do Número de Alunos.

Performance/Módulo	Módulo 1	Módulo 2	Módulo 3
Baixa	21	16	9
Intermediária	3	7	5
Alta	10	11	20

Fonte: Elaborada pelo autor (2026).

Ademais, para demonstrar a situação dos alunos em cada competência é favorável a apresentação das médias das competências em cada perfil por módulo, que pode ser identificado na tabela 5.

Tabela 5 – Tabela da Média das Competências em Relação aos Perfis por Módulo.

	Performance	Resiliência	Autonomia	Cooperação	Metacognição	Transdisciplinaridade	Flexibilidade	Criatividade
Módulo 1	Baixa	0.05	0.14	0.06	0.04	0.05	0.08	0.20
	Intermediária	0.00	0.41	0.35	0.54	0.06	0.00	0.39
	Alta	1.00	0.54	0.34	0.15	0.27	0.52	0.48
Módulo 2	Baixa	0.12	0.22	0.08	0.06	0.06	0.22	0.28
	Intermediária	0.57	0.42	0.43	0.46	0.13	0.45	0.25
	Alta	0.73	0.43	0.35	0.13	0.22	0.35	0.55
Módulo 3	Baixa	0.11	0.14	0.11	0.11	0.08	0.15	0.46
	Intermediária	0.60	0.28	0.37	0.44	0.14	0.51	0.56
	Alta	0.90	0.44	0.27	0.16	0.18	0.34	0.34

Fonte: Elaborada pelo autor (2026).

Cabe novamente esclarecer que os nomes dos perfis apresentados são para maior interpretabilidade pedagógica, mas é visível que retratam alunos que estariam abaixo, no meio, e acima da média. Um exemplo disso seriam as notas médias dos grupos de Baixa Performance onde dificilmente há competências nas quais passam dos 0.2.

4.3 Avaliação do Sistema de Recomendação

Após a contextualização da dinâmica de recomendação e a caracterização dos agrupamentos referentes aos perfis de estudantes, esta secção apresenta a avaliação quantitativa e pedagógica do sistema proposto. A análise fundamenta-se na utilização conjugada de duas frentes: (1) a aplicação da métrica *Average Normalized Gain* (ANG) para a avaliação do ganho pedagógico e à mobilidade dos alunos entre os *clusters*, confrontando o nível de adesão (aderência) às sugestões do modelo preditivo; e (2) a análise das correlações entre as atividades executadas e as competências desenvolvidas, com o intuito de identificar efeitos indiretos.

4.3.1 Avaliação por Utilização de Métricas entre *Clusters* e *Average Normalized Gain* (ANG) Sobre a Aderência

A avaliação de um Sistema de Recomendação Educacional (SRE) difere substancialmente da métrica de conversão utilizada em sistemas comerciais. No contexto deste projeto, é imperativo provar que a referida recomendação resultou numa evolução cognitiva. Para capturar essa dinâmica, a avaliação foi estruturada usando a transição dos alunos entre os *clusters* em comparativo com a aderência, cruzando o comportamento do aluno com o seu ganho de aprendizagem dentro do seu respetivo *cluster*.

Primeiramente, a base de estudantes foi segregada em dois subgrupos comportamentais durante os módulos 2 e 3:

- **Aderentes:** Estudantes que seguiram de forma consistente as atividades personalizadas recomendadas (> 0.5);
- **Não Aderentes:** Estudantes que não seguiram de forma consistente as atividades personalizadas recomendadas (< 0.5).

Para sistematizar a observação do impacto da aderência na trajetória académica ao longo das transições entre os módulos, estabeleceram-se três métricas comportamentais e seus respectivos **exemplos**:

- **Taxa de Ascensão:** Ocorre quando o estudante inicia num *cluster* de desempenho intermediário e, ao longo dos módulos, consegue migrar para o *cluster* de alta performance.
- **Taxa de Resgate:** Caracteriza-se quando o aluno inicia em zona de baixa performance e consegue recuperar o seu desempenho, indo para *clusters* superiores nas instâncias seguintes graças ao sistema de recomendação.

- **Taxa de Queda:** Mede a taxa de queda entre *clusters*, retratando alunos que estavam em um *cluster* superior e não conseguiram manter as notas nas competências.

Em complemento a esta análise de transição, para isolar o impacto da proficiência prévia do aluno (já que um aluno do *cluster* de “Alta Performance” tem menos margem para crescer do que um aluno do *cluster* de “Baixa Performance”), utilizou-se a métrica de Ganho Normalizado Médio (*Average Normalized Gain - g*), estabelecida por Hake (HAKE, 1998). Esta métrica calcula o ganho de aprendizagem efetivo em relação ao máximo ganho possível que o estudante ainda tinha disponível.

Matematicamente, o ganho normalizado (*g*) para as transições entre os módulos instrucionais é expresso pela equação 4.3:

$$g = \frac{\%S_{final} - \%S_{inicial}}{100\% - \%S_{inicial}} \quad (4.3)$$

Onde a variável $\%S_{inicial}$ representa a nota ou proficiência percentual do aluno no módulo base (ex: módulo 1) e $\%S_{final}$ corresponde ao desempenho alcançado no módulo subsequente (ex: módulo 3). O denominador atua como um fator de normalização.

Dito isso, com os grupos (*clusters*) estabelecidos e as métricas definidas foi elaborado o algoritmo para essas comparações.

Em primeiro lugar foi feito o comparativo das Taxas de: Ascensão, Resgate e Queda que relacionam os *cluster* aos alunos aderentes ou não, tais dados podem ser observados nas Tabelas 6 e 7.

Tabela 6 – Tabela das Taxas por Aderência do Módulo 1 para o Módulo 2.

Módulo 1 para Módulo 2			
	TA	TR	TQ
Não aderentes	0.143	0.143	0
Aderentes	0.259	0.429	0.154

Fonte: Elaborada pelo autor (2026).

Onde a Taxa de Ascensão é representado por TA, a Taxa de Resgate por TR e a Taxa de Queda por TQ. Assim, na tabela 6 é possível afirmar que: alunos aderentes as recomendações no módulo 1 atingiram uma camada superior nos *clusters* no módulo 2, e saíram do grupo de baixa performance, no entanto, houve uma pequena parcela que caiu para uma camada inferior. Já na tabela 7 representando do módulo 2 para o módulo 3, é possível observar novamente que alunos que foram aderentes conseguiram sair de grupos

Tabela 7 – Tabela das Taxas por Aderência do Módulo 2 para o Módulo 3.

Módulo 2 para Módulo 3			
	TA	TR	TQ
Não aderentes	0.154	0.143	0.167
Aderentes	0.524	0.778	0

Fonte: Elaborada pelo autor (2026).

em camadas inferiores para superiores e diferentemente da tabela 6, não houve queda dos alunos aderentes, somente dos alunos não aderentes.

Outro comparativo feito foi a relação entre o ANG e a aderência dos alunos, por meios empíricos e para uma melhor didática a porcentagem de aderência foi dividida da seguinte forma: Baixa, menor de 33%, Média, entre 33% e 66%, Alta, acima de 66%. Tais resultados são possíveis de visualizar nas tabelas 8 e 9.

Tabela 8 – Tabela de Ganho Normalizado Médio (ANG) por Aderência do Módulo 1 para o Módulo 2.

Ganho Normalizado Médio por Nível de Aderência	
Baixa (< 33%):	+0.0720
Média (33-66%):	-0.0126
Alta (> 66%):	+0.0528

Fonte: Elaborada pelo autor (2026).

Tabela 9 – Tabela de Ganho Normalizado Médio (ANG) por Aderência do Módulo 2 para o Módulo 3.

Ganho Normalizado Médio por Nível de Aderência	
Baixa (< 33%):	+0.0017
Média (33-66%):	+0.0961
Alta (> 66%):	+0.0772

Fonte: Elaborada pelo autor (2026).

Ao observar a tabela 8 o ANG foi maior e para os alunos de baixa aderência seguido dos alunos de alta aderência, no entanto, para os alunos de aderência média se obteve um pequeno decaimento. Entretanto, ao observar a tabela 9 pode-se visualizar um maior

efeito do sistema de recomendação, em que alunos de média e alta aderência tiveram um ANG maior enquanto os alunos de baixa aderência obtiveram uma melhora minúscula em relação aos outros.

Ao observar os resultados tanto da comparação das taxas de ascensão, de resgate e queda com a aderência, quanto os resultados obtidos pela comparação do ANG fica evidente a sugestão da aderência as recomendações influenciar de forma positiva os alunos. Entre os principais resultados podemos citar:

- A diferença expressiva entre TA em comparação aos alunos aderentes e não aderentes, com aproximadamente 10% de variação do módulo 1 para o módulo 2 e de aproximadamente 37% do módulo 2 para o módulo 3
- A diferença expressiva entre TR em comparação aos alunos aderentes e não aderentes, com aproximadamente 28% de variação do módulo 1 para o Módulo 2 e de aproximadamente 64% do módulo 2 para o módulo 3
- O ANG referente a aderência no módulo 1 para o módulo 2 e do módulo 2 para o módulo 3, em que inicialmente alunos mais aderentes a realização das atividades propostas pelas recomendações não foram diferenciados daqueles que tiveram baixa aderência, no entanto, na fase final da recomendação foi possível interpretar claramente esta distinção com média de ganho chegando quase a 10% para alunos de aderência média e quase a 8% para alunos de aderência alta.

De forma geral percebemos que do módulo 1 ao módulo 2, o sistema de recomendação não foi tão efetivo comparado aos resultados do módulo 2 ao módulo 3, talvez por um processo de adaptação dos alunos ao sistema de recomendação proposto.

4.3.2 Avaliação das Correlações das Atividades e Competências entre *Clusters*

No documento de [Oliveira \(2026\)](#), defende-se que as atividades pedagógicas do Moodle não atuam de forma isolada sobre uma única habilidade, mas apresentam correlações com múltiplas competências do sujeito complexo (tais como Autonomia, Cooperação, Resiliência e Transdisciplinaridade). Sob essa perspectiva, uma atividade rotulada prioritariamente para o desenvolvimento de uma competência primária pode, simultaneamente, interferir sobre outras competências correlatas.

Diante dessa premissa, foi realizada a avaliação empírica de tais correlações, buscando quantificar a intensidade e a direção dos impactos gerados pelas atividades do AVA nas competências correlatas.

Antes de mais nada, é necessário conhecer as correlações ali presentes, tais dados foram abordados no documento de [Oliveira \(2026\)](#) em que foram estabelecidas as correlações. Tais dados podem ser observados na tabela 10.

Tabela 10 – Tabela de Correlações.

Atividade	Rotulos com Correlação Significante		%
Glossário	Autonomia	Metacognição	44,70%
Base de Dados	Autonomia	Criatividade	43,30%
Chat	Cooperação	Flexibilidade	43,30%
Escolha	Autonomia	Metacognição	33,30%
Fórum	Cooperação	Transdisciplinaridade	63,30%
Workshop	Autonomia	Metacognição	53,30%
Lição	Autonomia	Resiliência	40,00%
Wiki	Autonomia	Cooperação	56,70%

Fonte: Adaptado de

([OLIVEIRA, 2026](#)).

Além disso, para essa análise, deve-se conhecer as atividades realizadas pelos alunos nos módulos 2 e 3. Para cada atividade que o aluno realizou, foi aplicado um procedimento computacional visando cruzar o histórico de proficiência com a matriz de correlações.

A operacionalização deste cálculo seguiu uma lógica fundamentada na variação de notas (Delta, Δ) e no peso das correlações na Tabela 10. O procedimento foi dividido nas seguintes etapas:

1. Mapeamento de Execuções: Primeiramente, identificou-se quais atividades (como Fórum, Wiki, *Quiz*, Lição, entre outras) foram efetivamente executadas por cada aluno durante as transições de aprendizagem.

2. Cálculo do Efeito Indireto: Para cada atividade realizada que possuía uma correlação mapeada na Tabela 10, o algoritmo calculou a variação bruta da nota do aluno na competência afetada (Δc) entre o módulo atual e o anterior. O impacto real da atividade, denominado Efeito Indireto, foi obtido multiplicando-se essa variação bruta pelo peso da correlação correspondente (W_{corr}), conforme formalizado pela Equação 4.4:

$$Efeito_Indireto = \Delta c \times W_{corr} \quad (4.4)$$

Nesta equação, Δc representa a diferença exata da proficiência na competência correlata, enquanto W_{corr} é o peso extraído da matriz teórica. Por exemplo, se um aluno executa um “Fórum”, avalia-se a variação bruta de sua nota em “Cooperação” multiplicada pelo peso de 0,633.

3. Agrupamento por *Clusters* e Consolidação: Para mitigar o efeito de valores atípicos (*outliers*), os dados foram agrupados considerando três dimensões: o *cluster* de origem do aluno, a atividade realizada e a competência afetada.

Para cada um desses cruzamentos, calculou-se a média do Delta Bruto, a média do Efeito Indireto e contabilizou-se o número total de ocorrências (quantas vezes a atividade foi executada naquele *cluster*).

Para facilitar a visualização e a interpretação pedagógica, os dados consolidados do módulo 1 para o módulo 2 estão expostos na Tabela 11.

Tabela 11 – Resultados das Correlações do Módulo 1 para o Módulo 2.

Cluster	Atividade	Competência Afetada	Efeito Indireto	Δ Bruto	Execuções
Alta Perf.	Base de Dados	Criatividade	+0.0014	+0.0033	6
	Base de Dados	Autonomia	-0.0375	-0.0867	6
	Chat	Cooperação	+0.0130	+0.0300	2
	Chat	Flexibilidade	-0.0238	-0.0550	2
	Escolha/Quiz	Metacognição	+0.0583	+0.1750	4
	Escolha/Quiz	Autonomia	-0.0433	-0.1300	4
	Fórum	Cooperação	+0.0269	+0.0425	4
	Fórum	Transdisciplinaridade	-0.0285	-0.0450	4
	Glossário	Metacognição	+0.0330	+0.0737	8
	Glossário	Autonomia	-0.0570	-0.1275	8
	Lição	Autonomia	+0.0020	+0.0050	4
	Lição	Resiliência	-0.4000	-1.0000	4
Baixa Perf.	Base de Dados	Autonomia	+0.0783	+0.1808	13
	Base de Dados	Criatividade	+0.0396	+0.0915	13
	Chat	Flexibilidade	+0.3594	+0.8300	1
	Chat	Cooperação	+0.0563	+0.1300	1
	Escolha/Quiz	Autonomia	+0.1004	+0.3015	13
	Escolha/Quiz	Metacognição	+0.0392	+0.1177	13
	Fórum	Cooperação	+0.0643	+0.1015	13
	Fórum	Transdisciplinaridade	+0.0370	+0.0585	13
	Glossário	Autonomia	+0.0808	+0.1808	13
	Glossário	Metacognição	+0.0461	+0.1031	13
	Lição	Autonomia	+0.1535	+0.3837	8
	Lição	Resiliência	0.0000	0.0000	8
Intermediário	Base de Dados	Autonomia	-0.1364	-0.3150	2
	Base de Dados	Criatividade	-0.1624	-0.3750	2
	Escolha/Quiz	Autonomia	-0.0416	-0.1250	2
	Escolha/Quiz	Metacognição	-0.1166	-0.3500	2
	Fórum	Transdisciplinaridade	+0.2089	+0.3300	1
	Fórum	Cooperação	+0.1393	+0.2200	1
	Glossário	Autonomia	-0.0969	-0.2167	3
	Glossário	Metacognição	-0.2190	-0.4900	3

Fonte: Elaborada pelo autor (2026).

De forma complementar, a Tabela 12 sintetiza a evolução cognitiva na transição do módulo 2 para o módulo 3, permitindo a observação do efeito às recomendações do sistema.

Tabela 12 – Resultados das Correlações do Módulo 2 para o Módulo 3.

Cluster	Atividade	Competência Afetada	Efeito Indireto	Δ Bruto	Execuções
Alta Perf.	Base de Dados	Autonomia	-0.0411	-0.0950	10
	Base de Dados	Criatividade	-0.1377	-0.3180	10
	Chat	Flexibilidade	-0.0179	-0.0413	8
	Chat	Cooperação	-0.0216	-0.0500	8
	Escolha/Quiz	Metacognição	+0.0311	+0.0933	6
	Escolha/Quiz	Autonomia	-0.0477	-0.1433	6
	Fórum	Cooperação	-0.0316	-0.0500	9
	Fórum	Transdisciplinaridade	-0.0450	-0.0711	9
	Glossário	Metacognição	+0.0260	+0.0582	11
	Glossário	Autonomia	-0.0504	-0.1127	11
	Lição	Resiliência	+0.1091	+0.2727	11
	Lição	Autonomia	-0.0451	-0.1127	11
	Wiki	Cooperação	-0.0249	-0.0440	10
	Wiki	Autonomia	-0.0522	-0.0920	10
Baixa Perf.	Base de Dados	Autonomia	+0.0273	+0.0631	13
	Base de Dados	Criatividade	-0.0603	-0.1392	13
	Chat	Cooperação	+0.0455	+0.1050	12
	Chat	Flexibilidade	+0.0119	+0.0275	12
	Escolha/Quiz	Autonomia	+0.0033	+0.0100	8
	Escolha/Quiz	Metacognição	+0.0008	+0.0025	8
	Fórum	Cooperação	+0.0515	+0.0814	14
	Fórum	Transdisciplinaridade	+0.0330	+0.0521	14
	Glossário	Autonomia	+0.0217	+0.0486	14
	Glossário	Metacognição	-0.0089	-0.0200	14
	Lição	Resiliência	+0.1333	+0.3333	12
	Lição	Autonomia	+0.0190	+0.0475	12
	Wiki	Cooperação	+0.0582	+0.1027	15
	Wiki	Autonomia	+0.0412	+0.0727	15
Intermediário	Base de Dados	Autonomia	-0.0450	-0.1040	5
	Base de Dados	Criatividade	-0.0450	-0.1040	5
	Chat	Cooperação	-0.0815	-0.1883	6
	Chat	Flexibilidade	-0.1335	-0.3083	6
	Escolha/Quiz	Autonomia	-0.0766	-0.2300	3
	Escolha/Quiz	Metacognição	-0.0977	-0.2933	3
	Fórum	Transdisciplinaridade	-0.0253	-0.0400	7
	Fórum	Cooperação	-0.1022	-0.1614	7
	Glossário	Autonomia	-0.0268	-0.0600	7
	Glossário	Metacognição	-0.1092	-0.2443	7
	Lição	Resiliência	+0.0800	+0.2000	5
	Lição	Autonomia	-0.0688	-0.1720	5
	Wiki	Autonomia	-0.0340	-0.0600	7
	Wiki	Cooperação	-0.0915	-0.1614	7

Fonte: Elaborada pelo autor (2026).

Ao analisar as Tabelas 11 e 12 é possível interpretar três constatações empíricas para a avaliação do sistema de recomendação educacional proposto:

1. Efeito Teto (*Ceiling Effect*)

Os dados do *cluster* de Alta Performance comportam-se de maneira condizente com o fenômeno estatístico conhecido como Efeito Teto (*Ceiling Effect*). Na literatura de avalia-

ção educacional, este efeito ocorre quando os sujeitos analisados já se encontram próximos ao limite máximo da escala de medição, o que restringe matematicamente a possibilidade de registrar ganhos absolutos significativos (HAKE, 1998). Observa-se no grupo de Alta Performance uma predominância de variações brutas (Δ) negativas ou extremamente próximas a zero (como o ganho de +0.0014 em Criatividade no módulo 1 para o módulo 2). Como estes alunos já iniciam a disciplina com o domínio cognitivo quase total, o cálculo do impacto transversal torna-se restrito. Nesses casos, conforme fundamenta a métrica de Average Normalized Gain proposta por Hake (1998), as recomendações para este grupo atuam mais na manutenção do engajamento e na prevenção de quedas de rendimento do que no salto expressivo de notas.

2. Amadurecimento dos Alunos em Relação à Atividades

O cruzamento do desempenho da ferramenta “Lição” na competência de “Resiliência” no *cluster* de Baixa Performance revela o amadurecimento em relação à atividade. Na primeira recomendação, a execução da atividade de lição gerou um ganho absoluto nulo (Impacto = 0.0000 em 8 execuções). Contudo, a exposição a mesma atividade durante a recomendação do módulo 2 para o módulo 3 resultou no maior pico transversal do período para este grupo (Impacto = +0.1333, com Δ = +0.3333 em 12 execuções). Isso pode comprovar que competências, como a Resiliência, não respondem de forma imediata a estímulos de curto prazo, demandando maturação e consistência instrucional.

3. Eficácia de Resgate de Perfil de Baixa Performance

Um dos pontos principais a serem destacados se encontra no comportamento do *cluster* de Baixa Performance. Ao contrário dos demais grupos, estes alunos registraram impactos indiretos majoritariamente positivos em todas as ferramentas. Destacam-se as 13 execuções de “Fórum” no primeiro ciclo, que impulsionaram simultaneamente a Cooperação (+0.0643) e a Transdisciplinaridade (+0.0370). Fica evidente que as recomendações do sistema possuem forte ligação com mecanismo de “Resgate”, demonstrando que, para discentes com lacunas ativas, a personalização de atividades do Moodle reflete em ganhos mensuráveis nas competências.

4.3.3 Estudo de Casos: Análise Individualizada de Aderência e Correlações

A fim de proporcionar uma compreensão qualitativa e aprofundada sobre a atuação prática do sistema de recomendação, esta subseção analisa dois casos individuais extraídos da base de dados empírica. O estudo de casos permite detalhar como as taxas de adesão (aderência) influenciaram diretamente as métricas de Resgate, Ascensão e Queda, bem como evidenciar os impactos previstos pela matriz teórica de correlações.

Caso 1: Resgate e Ascensão via Crescimento de Aderência (Aluno ID 60)

O Aluno 60 iniciou o período letivo com um desempenho mediano-baixo, registrando

uma performance de 0.63 no módulo 1. Ao longo das transições, o estudante demonstrou uma curva crescente de engajamento: apresentou uma aderência de 57% às atividades recomendadas no módulo 2 e atingiu 100% de aderência no módulo 3.

Essa adesão massiva às recomendações refletiu-se em um cenário perfeito de Ascensão. No módulo 2, sua performance geral subiu para 0.88, culminando na nota máxima (1.00) ao final do módulo 3.

Sobre as correlações propostas neste estudo, o impacto da adesão é inquestionável. No módulo 1, o aluno possuía proficiência nula (0.00) na competência de *Metacognição*. Ao seguir as recomendações do DKVMN e engajar em atividades com forte peso para essa competência (como *Glossário* e *Escolha/Quiz*), sua nota de Metacognição subiu para 0.07 no módulo 2 e deu um salto expressivo para 0.51 no módulo 3. Este caso comprova que a execução das recomendações ao longo dos módulos de forma ativa (alta aderência) é capaz de preencher lacunas cognitivas específicas com sucesso.

Caso 2: O Efeito Teto e a Prevenção de Quedas na Alta Performance (Aluno ID 165)

O terceiro caso debruça-se sobre um discente de excelência (Aluno 165). Com desempenho elevado desde o início (0.88 de performance no módulo 1), este estudante registrou uma alta aderência às recomendações (100% de taxa de adesão nas atividades sugeridas nos módulos subsequentes).

Como previsto pelas limitações matemáticas discutidas nas seções anteriores, o Efeito Teto limitou a percepção de ganhos deste discente. Por exemplo, sua competência de *Autonomia* iniciou muito alta (0.62 no módulo 1). Apesar de realizar as atividades sugeridas, a nota de Autonomia não saltou; pelo contrário, ajustou-se para 0.58 no módulo 2 e 0.52 no módulo 3, gerando deltas negativos na matriz de efeitos indiretos.

Contudo, a aderência contínua não foi inútil: ela manifestou-se na métrica de Taxa de Queda. O cumprimento das atividades blindou o aluno cognitivamente contra a transição para um *cluster* de baixa performance ao final da disciplina, permitindo que sua performance geral atingisse a marca de 1.00 no módulo 3. Isso consolida a tese de que, para o *cluster* de Alta Performance, o sistema atua na manutenção do engajamento do aluno, cumprindo com o dever de segurar a performance do aluno e suas competências.

4.4 Sugestões para Adaptação do Sistema de Recomendação Dependente dos Resultados da Avaliação

A partir dos resultados quantitativos, estatísticos e pedagógicos obtidos ao longo do processo avaliativo, torna-se possível propor adaptações estratégicas para o aprimoramento da arquitetura do sistema de recomendação. Tais sugestões visam otimizar o

algoritmo preditivo e garantir maior impacto no desenvolvimento cognitivo dos estudantes.

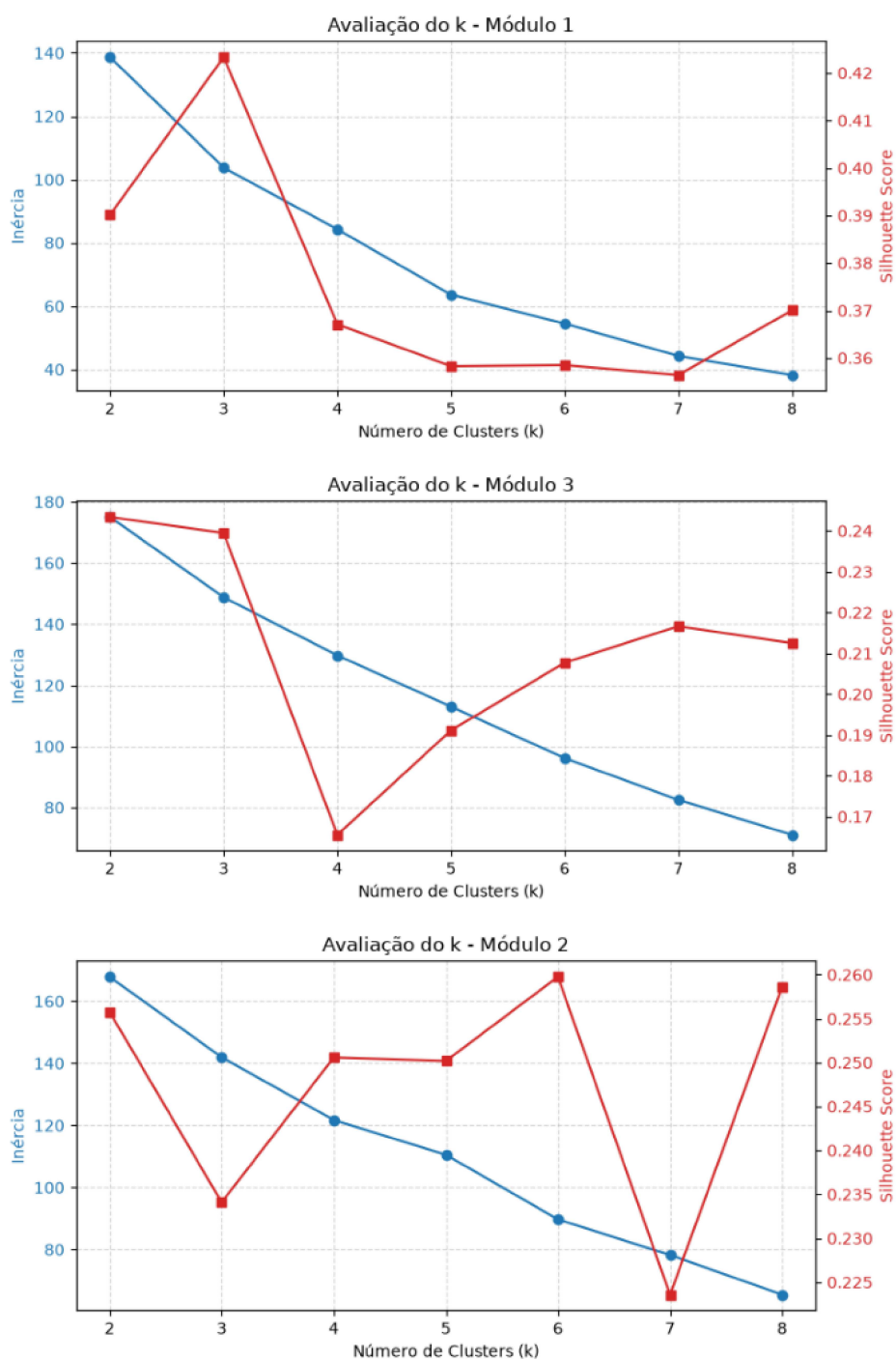
1. Expansão Longitudinal dos Módulos Adaptativos

Os achados empíricos demonstraram que o impacto do sistema de recomendação varia substancialmente de acordo com o tempo de exposição do discente às atividades. Em geral, através das taxas (TA, TR e TQ) e do ANG foi possível visualizar uma diferença na variação dos alunos aderentes ao sistema, no entanto, é possível visualizar, que do módulo 2 para o módulo 3, foram obtidas variações mais significativas em ambos os aspectos. Esse comportamento pode indicar que o desenvolvimento de competências complexas e ligadas a performance do aluno podem requerer mais ciclos. Recomenda-se, portanto, a ampliação do ciclo de aplicação para além de três módulos, permitindo acompanhar a consolidação contínua do aprendizado em trajetórias de longo prazo.

2. Recalibração das Pontuações Iniciais e Diagnóstico de Competências

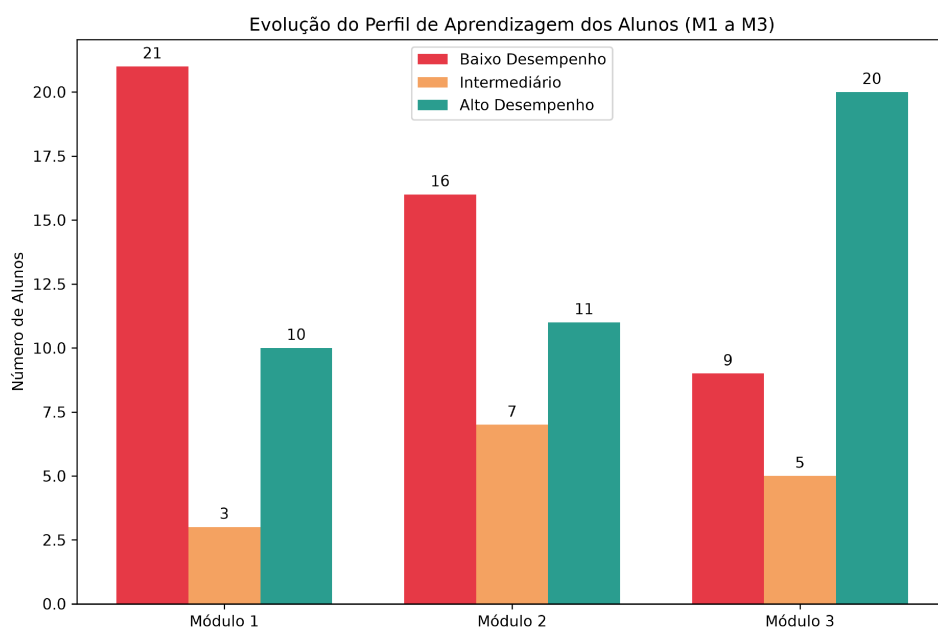
Observou-se que determinadas competências (como *Metacognição* e *Transdisciplinaridade*) registraram notas cronicamente reduzidas em quase a totalidade da amostra, independentemente do *cluster* de origem do discente e olhando caso a caso nos dados obtidos pelo sistema de recomendação se nota raras exceções com notas medianas. Essa estagnação em patamares baixos sugere a necessidade de recalibrar a escala inicial de proficiência atribuída pelo modelo DKVMN ou reavaliar o grau de exigência das atividades rotuladas para essas habilidades. Ajustar os limiares de avaliação inicial prevenirá diagnósticos de severidade excessiva e permitirá medir variações de proficiência de forma mais sensível.

Figura 2 – Gráficos por Módulo com *Silhouette Score* e suas variações (Inércia) em função de k .



Fonte: Elaborada pelo autor (2026).

Figura 3 – Gráfico de Performance x Módulo em Função do Número de Alunos.



Fonte: Elaborada pelo autor (2026).

5 Conclusão

O presente trabalho teve como objetivo central avaliar a eficácia do Sistema de Recomendação Educacional (SRE) integrado ao Ambiente Virtual de Aprendizagem Moodle, suportado pelo modelo preditivo *Dynamic Key-Value Memory Networks* (DKVMN). Esta investigação procurou avaliar o impacto pedagógico de recomendações personalizadas no desenvolvimento de competências complexas dos estudantes, ancoradas na teoria do sujeito complexo (MORIN, 2000).

A metodologia adotada para esta avaliação baseou-se em diferentes meios de avaliação, seja por métricas existentes ou análises decorrentes de processos de agrupamento. Através da aplicação do algoritmo de agrupamento K-means, foi possível segmentar os estudantes em três perfis cognitivos distintos: Baixa Performance, Intermediário e Alta Performance. Esta estratificação permitiu uma análise contextualizada, rejeitando a premissa de que uma única abordagem avaliativa ou métrica se adequaria de forma justa a toda a turma. Em complemento, a introdução de métricas de avaliação educacional, como o *Average Normalized Gain* (ANG) e as taxas comportamentais (Taxa de Ascensão, Resgate e Queda), forneceram os instrumentos necessários para quantificar de forma objetiva a evolução cognitiva dos estudantes ao longo dos três módulos de avaliação, possuindo clara distinção do desempenho acadêmico.

Os resultados empíricos obtidos comprovaram que a adesão ativa (aderência) às recomendações geradas pelo sistema é um fator determinante para o sucesso pedagógico. Foi possível constatar que os estudantes com alta taxa de adesão apresentaram níveis significativamente superiores de Ganho Normalizado Médio, especialmente na fase de consolidação correspondente à transição do módulo 2 para o módulo 3. Mais do que um mero aumento nas classificações, o sistema provou ser uma ferramenta favorável no contexto educacional: as Taxas de Resgate evidenciaram que estudantes alocados inicialmente no perfil de Baixa Performance conseguiram, através das atividades recomendadas, suprir as suas lacunas e ascender a perfis de melhor desempenho. Atividades iterativas, como o “Fórum” e a “Lição”, desempenharam um papel crucial na maturação de competências cruzadas, gerando impactos indiretos altamente positivos em áreas como a cooperação, a resiliência e a transdisciplinaridade.

Por outro lado, a análise empírica confirmou a ocorrência do Efeito Teto (*Ceiling Effect*) no *cluster* de Alta Performance. Como teorizado na literatura, os estudantes que já se encontravam no limiar superior de excelência no início da disciplina apresentaram margens matemáticas reduzidas para a progressão de notas brutas. Contudo, o sistema não se revelou ineficaz para este grupo; pelo contrário, a adesão às sugestões atuou como

um mecanismo de prevenção de declínio cognitivo, traduzindo-se numa Taxa de Queda nula no último período avaliativo e na manutenção contínua do engajamento estudantil face aos desafios da plataforma.

Apesar dos resultados positivos que sugerem a validação a tese inicial, este estudo permitiu também identificar oportunidades de aprimoramento para o sistema de recomendação proposto. O comportamento de progresso contínuo dos dados sugere que a consolidação de competências mais complexas demanda uma exposição superior a três ciclos de aprendizagem, justificando a necessidade da implementação de uma maior quantidade de módulos. Adicionalmente, verificou-se a necessidade de uma recalibração dos limiares iniciais de avaliação do modelo DKVMN, em particular para competências como a Metacognição, por forma a evitar diagnósticos iniciais de severidade excessiva e refinar ainda mais a precisão das futuras recomendações.

Como propostas para trabalhos futuros, recomenda-se a aplicação e avaliação deste sistema de recomendação em amostras mais amplas e em diferentes contextos disciplinares, a fim de verificar a generalização dos resultados. Sabendo que a adesão dos alunos às atividades é o pilar para a evolução cognitiva, a investigação e integração de mecanismos ativos para estimular a aderência apresenta-se promissor, talvez atividades que tem um maior apelo visual seriam vantajosas. Com uma maior quantidade de amostras e obtenção de uma maior aderência às recomendações, é possível realizar novamente uma avaliação sobre os dados obtidos, a fim de testar o sistema de recomendação com uma base mais sólida e buscar novos métodos para avaliação, visto que, a pequena quantidade de amostras se torna ineficiente para concluir alguns aspectos relacionados a atividades e competências.

Em suma, este projeto sugere a validação da eficácia dos Sistemas de Recomendação Educacional baseados em modelos avançados de *Knowledge Tracing*. A ponte estabelecida entre as técnicas de Inteligência Artificial não supervisionada e as teorias pedagógicas demonstra que a personalização das atividades interagem positivamente para o aprendizado dos estudantes.

Referências

- AKHADAM, A.; KBIBCHI, O.; MEKOUAR, L.; IRAQI, Y. A comparative evaluation of recommender systems tools. **IEEE Access**, 2025. Citado 3 vezes nas páginas 13, 20 e 21.
- ALDOWAH, H.; AL-SAMARRAIE, H.; FAUZY, W. M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. **Telematics and Informatics**, Elsevier, v. 37, p. 13–49, 2019. Citado na página 17.
- APINO, E.; RETNAWATI, H. Developing instructional design to improve mathematical higher order thinking skills of students. **Journal of Physics: Conference Series**, IOP Publishing, 2017. Citado na página 12.
- CABINESS, C.; DONOVAN, L.; GREEN, T. D. Integrating wikis in the support and practice of historical analysis skills. **TechTrends**, Springer, 2013. Citado na página 12.
- CHURCHES, A. **Bloom’s Digital Taxonomy**. 2008. 1-6 p. Citado na página 19.
- DUTT, A.; ISMAIL, M. A.; HERAWAN, T. A systematic review on educational data mining. **IEEE Access**, IEEE, v. 5, p. 15991–16005, 2017. Citado na página 17.
- ESTER, M.; KRIEGEL, H.-P.; SANDER, J.; XU, X. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: **Kdd**. Portland, Oregon: AAAI Press, 1996. v. 96, n. 34, p. 226–231. Citado na página 47.
- HAKE, R. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. **American Journal of Physics**, 1998. Citado 5 vezes nas páginas 13, 17, 20, 31 e 37.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques**. San Francisco: Morgan kaufmann, 2011. Citado na página 27.
- HERLOCKER, J. L.; KONSTAN, J. A.; TERVEEN, L. G.; RIEDL, J. Evaluating collaborative filtering recommender systems. **ACM Transactions on Information Systems (TOIS)**, v. 22, n. 1, p. 5–53, 2004. Citado na página 17.
- JAIN, A. K. Data clustering: 50 years beyond k-means. **Pattern Recognition Letters**, Elsevier, v. 31, n. 8, p. 651–666, 2010. Citado 4 vezes nas páginas 13, 16, 22 e 27.
- KLAŠNJA-MILIĆEVIĆ, A.; IVANOVIĆ, M.; NANOPOULOS, A. Recommender systems in e-learning environments: a survey of the state-of-the-art and possible extensions. **Artificial Intelligence Review**, Springer, v. 44, n. 4, p. 571–604, 2015. Citado na página 15.
- KODINARIYA, T. M.; MAKWANA, P. R. Review on determining number of cluster in k-means clustering. **International Journal of Advance Research in Computer Science and Management Studies**, v. 1, n. 6, p. 90–95, 2013. Citado na página 28.
- MORIN, E. **Introdução ao Pensamento Complexo**. 1. ed. Porto Alegre: Sulina, 2000. Citado 4 vezes nas páginas 12, 19, 25 e 42.

OLIVEIRA, G. P. **Método de Recomendação Personalizada de Atividades em Educação a Distância: Uma Abordagem Baseada em Pensamento Complexo e Inteligência Artificial**. Tese (Qualificação de Doutorado) — Universidade Federal de Uberlândia, 2026. Orientadora: Prof^ª Dra. Márcia Aparecida Fernandes; Coorientadora: Prof^ª Dra. Jaqueline Maissiat. Citado 7 vezes nas páginas [12](#), [13](#), [19](#), [25](#), [26](#), [33](#) e [34](#).

OLIVEIRA, G. P.; FERNANDES, M. A.; MAISSIAT, J. Correlacionando habilidades cognitivas às atividades do moodle à luz da teoria da complexidade. **Anais do Simpósio Brasileiro de Informática na Educação, (SBIE)**, 2022. Citado na página [12](#).

_____. **Validating the process for labeling Moodle activities according to Complex Thinking Theory**. 2023. Citado na página [12](#).

PACHECO, C. S.; HERRERA, C. I. A conceptual proposal and operational definitions of the cognitive processes of complex thinking. **Science Direct**, 2021. Citado na página [19](#).

RIVERA, A. C.; TAPIA-LEON, M.; LUJAN-MORA, S. Recommendation systems in education: A systematic mapping study. **International Conference on Information Technology & Systems (ICITS)**, Springer, 2018. Citado na página [15](#).

ROMERO, C.; VENTURA, S. Educational data mining: a review of the state of the art. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, IEEE, v. 40, n. 6, p. 601–618, 2010. Citado 2 vezes nas páginas [17](#) e [26](#).

ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of computational and applied mathematics**, Elsevier, v. 20, p. 53–65, 1987. Citado na página [28](#).

ZHANG, J.; SHI, X.; KING, I.; YEUNG, D.-Y. Dynamic key-value memory networks for knowledge tracing. In: **Proceedings of the 26th International Conference on World Wide Web**. Genebra: International World Wide Web Conferences Steering Committee, 2017. p. 765–774. Citado 4 vezes nas páginas [13](#), [16](#), [17](#) e [22](#).

Apêndices

APÊNDICE A – Exploração de Algoritmos Baseados em Densidade

Embora a metodologia principal deste trabalho tenha adotado o algoritmo K-Means para a segmentação dos perfis de aprendizagem (conforme discutido na Seção 4.2), é pertinente explorar alternativas de agrupamento que lidem com características distintas dos dados educacionais. Uma das principais limitações do K-Means é a necessidade de definição prévia do número de agrupamentos (k) e a sua elevada sensibilidade a valores atípicos (*outliers*), forçando a alocação de todos os estudantes a um centroide, mesmo aqueles com trajetórias de aprendizagem anômalas.

Nesse contexto, o algoritmo DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) desponta como uma alternativa robusta e complementar. Introduzido no trabalho de Ester et al. (1996), o DBSCAN afasta-se do conceito de centroides geométricos e define *clusters* como regiões contínuas de alta densidade de dados, separadas por regiões de baixa densidade.

A arquitetura do DBSCAN baseia-se em dois parâmetros fundamentais descritos por Ester et al. (1996):

- **Eps (ϵ):** O raio máximo da vizinhança de um ponto. Representa a distância máxima aceitável para que dois pontos sejam considerados pertencentes à mesma região densa.
- **MinPts:** O número mínimo de pontos (incluindo o próprio ponto) que devem estar contidos dentro do raio ϵ para que a região seja classificada como um *cluster*.

Com base nestes parâmetros, o algoritmo classifica os dados em três categorias: pontos centrais (*core points*), que possuem pelo menos MinPts dentro do seu raio; pontos de borda (*border points*), que não atingem o MinPts mas estão na vizinhança de um ponto central; e os pontos de ruído (*noise points*), que não pertencem a nenhuma das categorias anteriores (ESTER et al., 1996).

Aplicação no Contexto Educacional

A aplicação do DBSCAN em dados provenientes de Sistemas de Recomendação Educacional (SRE) apresenta um potencial analítico significativo. Diferente do K-Means, que divide a turma em grupos generalistas, o DBSCAN possui a capacidade de identificar

alunos que não se encaixam em nenhum padrão de aprendizagem tradicional, no contexto do algoritmo, são aqueles classificados algorítmicamente como ruído.

Pedagogicamente, um ponto de “ruído” pode representar um aluno com um perfil de competências altamente inconsistente, como um aluno que obteve nota máxima em Transdisciplinaridade, mas possui engajamento nulo em Autonomia e Cooperação, ou pode representar um aluno com notas máximas ou zeradas em todas as competências.

O isolamento destes perfis atípicos proporciona a oportunidade de realizar uma análise mais profunda e individualizada. Ao separar esses estudantes do restante da turma, torna-se possível investigar os motivos subjacentes a essas trajetórias não convencionais, compreendendo o porquê de tais situações ocorrerem. Isso permitiria para o sistema de recomendação, a intervenção diferenciada nesses casos.

Para fins de avaliação, o algoritmo DBSCAN foi executado sobre o conjunto de dados empíricos utilizado neste projeto. Contudo, observou-se que o tamanho reduzido da amostra limitou o potencial analítico da técnica. Embora o algoritmo tenha sido capaz de segregar indivíduos caracterizados como ruído, a pequena quantidade de dados tornou inviável a condução de uma investigação sobre esses perfis.

Do ponto de vista metodológico e científico, não é apropriado derivar conclusões pedagógicas do sistema com base nas anomalias de apenas um ou dois alunos isolados, podendo tirar conclusões precipitadas a respeito dos dados.

No entanto, estes testes preliminares indicam que a integração do DBSCAN possui grande potencial. Infere-se que, em trabalhos futuros que contemplem turmas de maior escala e um volume significativamente superior de amostras, a utilização de algoritmos baseados em densidade será plenamente justificável. Com uma base de dados expandida, a análise desse “ruído” educacional deixará de ser uma anomalia isolada para se tornar possivelmente um dos pontos centrais a serem analisados, permitindo mapear falhas no modelo instrucional e aprimorar a personalização das recomendações.