



Ministério da Educação
Secretaria de Educação Profissional e Tecnológica
Instituto Federal Catarinense
Campus Videira

THALES FERREIRA BATISTA

OTIMIZAÇÃO DE PLANOS DE ENSINO DE MATEMÁTICA

Videira
2025

THALES FERREIRA BATISTA

OTIMIZAÇÃO DE PLANOS DE ENSINO DE MATEMÁTICA

Projeto de Trabalho de Conclusão de Curso
apresentado ao Curso de graduação em Ciência da Computação do Instituto Federal Catarinense – Campus Videira para obtenção do título de bacharel em Ciência da Computação.

Orientador: Prof. Dr. Rafael Zanin (IFC)

Coorientador: Prof. Dr. Manassés Ribeiro (IFC)

Videira

2025

LISTA DE ILUSTRAÇÕES

Figura 1 – Fluxo PRISMA 2020 da revisão sistemática.	29
Figura 2 – Funil de seleção de estudos.	29
Figura 3 – Registros identificados por base de dados (inclui duplicatas) (n = 9.513). . .	30
Figura 4 – Distribuição de técnicas de IA nos estudos incluídos.	33
Figura 5 – Distribuição temporal dos estudos incluídos (2017-2026).	34
Figura 6 – Distribuição de scores de relevância dos estudos incluídos.	35

SUMÁRIO

1	INTRODUÇÃO	7
1.1	Problema de Pesquisa	7
1.2	Objetivos	8
1.2.1	Objetivo Geral	8
1.2.2	Objetivos Específicos	8
1.3	Justificativa	9
1.3.1	Relevância Acadêmica	9
1.3.2	Relevância Prática	10
1.4	Estrutura do Trabalho	10
1.4.1	Fase 1: Revisão Sistemática da Literatura (PTC)	10
1.4.2	Fase 2: Desenvolvimento do Protótipo (TCC)	11
1.4.3	Fase 3: Validação Experimental (TCC)	11
1.4.4	Organização do Documento	12
1.4.5	Delimitação de Escopo	12
2	FUNDAMENTAÇÃO TEÓRICA	13
2.1	Contexto e Relevância	13
2.2	Questões de Pesquisa	13
2.3	Técnicas Computacionais na Educação Matemática	14
2.3.1	Machine Learning e Inteligência Artificial	14
2.3.2	Learning Analytics e Educational Data Mining	14
2.3.3	Sistemas Tutores Inteligentes	15
2.3.4	Aprendizagem Adaptativa e Personalizada	16
2.4	Avaliação Automatizada e Métricas de Desempenho	16
2.4.1	Tipos de Avaliação	16
2.4.2	Métricas Computacionais	17
2.5	Justificativas Metodológicas para a Revisão Sistemática	17
2.5.1	Escolha de Múltiplas Bases de Dados	17
2.5.2	Critérios de Recorte Temporal	18
2.5.3	Abordagem Bilíngue	18
2.6	Contribuições Esperadas	18

3	METODOLOGIA	20
3.1	Protocolo da Revisão Sistemática	20
3.2	Estratégia de Busca	20
3.2.1	Bases de Dados e APIs	20
3.2.2	Estratégia de Busca Bilíngue	20
3.3	Critérios de Seleção (PICOS)	22
3.3.1	Critérios de Inclusão	22
3.3.2	Critérios de Exclusão	22
3.4	Processo de Seleção (PRISMA)	23
3.4.1	Identificação	23
3.4.2	Triagem	23
3.4.3	Elegibilidade	23
3.4.4	Inclusão	24
3.5	Deduplicação	24
3.6	Infraestrutura Tecnológica	24
3.6.1	Pipeline Automatizado	24
3.6.2	Reprodutibilidade	25
3.7	Limitações Metodológicas	25
4	RESULTADOS ESPERADOS	27
4.1	Fluxo PRISMA 2020	27
4.1.1	Identificação (n = 9.513)	27
4.1.2	Triagem (n = 6.948)	27
4.1.3	Elegibilidade (n = 1.945)	28
4.1.4	Incluídos (n = 16)	28
4.2	Estatísticas Descritivas	28
4.3	Síntese dos Estudos Incluídos	28
4.4	Análise Temática	32
4.4.1	Termos Mais Frequentes	32
4.4.2	Categorias Temáticas Emergentes	32
4.4.3	Distribuição por Abordagem Técnica e Finalidade	33
4.5	Resultados de Eficácia Reportados	33
4.5.1	Magnitude de Efeito	34

4.5.2	Métricas de Avaliação Empregadas	34
4.6	Limitações Identificadas	35
4.6.1	Limitações Técnicas	35
4.6.2	Limitações Pedagógicas	35
4.6.3	Limitações Metodológicas	36
4.6.4	Limitações Éticas	36
4.7	Mapeamento de Lacunas e Direcionamento para Fase 2	36
4.7.1	Lacunas Técnicas	36
4.7.2	Lacunas Pedagógicas	37
4.7.3	Direcionamento para Fase 2	37
4.8	Próximos Passos	37
4.8.1	Fase 2 — Desenvolvimento de Protótipo (fevereiro–julho/2026)	38
4.8.2	Fase 3 — Validação Experimental (julho–novembro/2026)	38
4.8.3	Síntese Final do TCC (outubro–novembro/2026)	38
5	CRONOGRAMA	39
	REFERÊNCIAS	40

1 INTRODUÇÃO

A educação matemática enfrenta o desafio constante de atender às diversas necessidades de aprendizagem dos alunos em salas de aula heterogêneas. A personalização do ensino, embora reconhecida como uma abordagem eficaz para melhorar o desempenho e o engajamento dos estudantes, é uma tarefa complexa e demorada para os professores (ZHANG, 2024; ZHANG; ZHU; FENG, 2025). A dificuldade em diagnosticar com precisão e em tempo hábil as competências e dificuldades individuais de cada aluno representa uma barreira significativa para a otimização dos planos de ensino (APPIAH-ODAME, 2024).

Nesse contexto, as tecnologias computacionais — como *machine learning*, análise de dados educacionais (*learning analytics*) e inteligência artificial — surgem como aliadas poderosas (USKOV et al., 2019; TJAHYADI, 2025). Essas técnicas oferecem o potencial de automatizar a avaliação diagnóstica e fornecer aos educadores informações estruturadas e acionáveis para intervenções pedagógicas mais direcionadas e eficazes (PEJIC; MOLCER; GULAČI, 2021; SOKKHEY et al., 2020). Estudos recentes demonstram que sistemas adaptativos baseados em dados podem melhorar significativamente o desempenho estudantil e reduzir o tempo necessário para alcançar objetivos de aprendizagem (JOSE et al., 2024).

1.1 PROBLEMA DE PESQUISA

O acompanhamento individualizado do progresso dos alunos em matemática é fundamental para um ensino de qualidade. No entanto, professores frequentemente carecem de ferramentas eficientes para diagnosticar as competências específicas de cada estudante em larga escala (KUMAR et al., 2022), o que dificulta a adaptação dos planos de ensino às necessidades reais da turma. A ausência de um sistema automatizado para essa finalidade leva a um ensino mais generalizado, que pode não atender adequadamente nem os alunos com dificuldades nem os mais avançados.

Apesar do crescente número de pesquisas sobre inteligência artificial aplicada à educação matemática, a literatura científica apresenta-se fragmentada e dispersa em múltiplas bases de dados, subdisciplinas e abordagens metodológicas. Pesquisadores e desenvolvedores enfrentam dificuldades para identificar quais técnicas computacionais são mais adequadas para diferentes objetivos pedagógicos e quais abordagens foram efetivamente validadas em contextos educacionais reais (MACLELLAN, 2017).

Diante desse contexto, o presente trabalho aborda o seguinte problema de pesquisa:

Como identificar e sintetizar, através de revisão sistemática da literatura, as principais técnicas computacionais aplicadas à educação matemática, de modo a fundamentar o desenvolvimento de uma ferramenta que auxilie professores na otimização de seus planos de ensino a partir da identificação automatizada das competências e dificuldades individuais dos alunos?

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Mapear e analisar sistematicamente as aplicações de técnicas computacionais — especialmente *machine learning*, *learning analytics* e sistemas de tutoria inteligente — no contexto da educação matemática, identificando tendências, lacunas de pesquisa e oportunidades para o desenvolvimento de ferramentas que apoiem o processo de ensino-aprendizagem personalizado.

1.2.2 Objetivos Específicos

- **OE1:** Realizar revisão sistemática da literatura seguindo o protocolo PRISMA 2020 para identificar estudos que apliquem técnicas computacionais no ensino da matemática, publicados nos últimos 10 anos (2015-2025).
- **OE2:** Identificar e categorizar as principais abordagens de inteligência artificial (Machine Learning, Deep Learning, NLP, Educational Data Mining) aplicadas à educação matemática reportadas na literatura científica.
- **OE3:** Classificar as aplicações identificadas segundo suas finalidades pedagógicas: tutoria inteligente, diagnóstico de dificuldades, avaliação automatizada, personalização de conteúdo, predição de desempenho e feedback adaptativo.
- **OE4:** Analisar criticamente as metodologias de avaliação utilizadas para validar a eficácia de sistemas computacionais em contextos educacionais, identificando boas práticas e limitações metodológicas.
- **OE5:** Mapear sistematicamente as lacunas de pesquisa, limitações técnicas e desafios reportados nos estudos incluídos, propondo direções prioritárias para o desenvolvimento de ferramentas computacionais eficazes em educação matemática.

- **OE6:** Criar um pipeline automatizado e reproduzível para coleta, processamento e análise de literatura científica, contribuindo para futuras revisões sistemáticas no campo e fundamentando o desenvolvimento do protótipo nas fases subsequentes do TCC.

1.3 JUSTIFICATIVA

A relevância deste projeto reside no seu potencial de impacto tanto para educadores quanto para alunos. Para os professores, a síntese sistemática das técnicas computacionais aplicadas à educação matemática oferece fundamentação científica sólida para a tomada de decisão sobre investimentos em tecnologia educacional e adoção de práticas baseadas em evidências (DEPREN; AŞKIN; ÖZ, 2017). Além disso, a identificação de abordagens eficazes pode orientar o desenvolvimento de ferramentas que reduzam a carga de trabalho manual de avaliação e forneçam *insights* acionáveis baseados em dados, permitindo que os educadores dediquem mais tempo à elaboração de estratégias pedagógicas personalizadas (MERTASARI; SASTRI; PASCIMA, 2023).

Para os alunos, um ensino adaptado às suas necessidades individuais — fundamentado em diagnósticos automatizados precisos — pode resultar em melhoria significativa na aprendizagem, na motivação e na redução da ansiedade em relação à matemática (JOSE et al., 2024). Estudos reportam ganhos de aprendizagem de 10 a 20% em sistemas adaptativos quando comparados a abordagens tradicionais, além de redução de até 20% no tempo necessário para atingir objetivos pedagógicos (ZHANG; ZHU; FENG, 2025).

1.3.1 Relevância Acadêmica

A realização de uma revisão sistemática sobre técnicas computacionais na educação matemática justifica-se por múltiplas razões acadêmicas:

1. **Consolidação do Conhecimento Disperso:** A literatura sobre IA educacional está fragmentada em conferências de Educação, Ciência da Computação, Inteligência Artificial e Tecnologia Educacional. Não existe consenso sobre terminologias e classificações das aplicações.
2. **Rigor Metodológico:** A metodologia PRISMA 2020 garante transparência, reprodutibilidade e qualidade na seleção de estudos. O uso de múltiplas bases de dados (Crossref, OpenAlex, Semantic Scholar, CORE) minimiza viés de publicação, e a aplicação de critérios de elegibilidade explícitos assegura qualidade dos estudos incluídos.

3. **Lacuna de Pesquisa Identificada:** Revisões anteriores não utilizaram pipelines automatizados e reproduzíveis. Poucos estudos focam especificamente na interseção entre IA e educação matemática, e falta análise crítica das metodologias de avaliação empregadas nos estudos primários (USKOV et al., 2019).

1.3.2 Relevância Prática

Do ponto de vista prático, este trabalho contribui para:

1. **Orientação para Desenvolvimento de Sistemas:** Desenvolvedores de software educacional necessitam evidências sobre quais técnicas são mais eficazes. A síntese da literatura fornece diretrizes baseadas em dados empíricos para escolha de abordagens técnicas (MACLELLAN, 2017).
2. **Suporte a Políticas Públicas:** Formuladores de políticas educacionais precisam de sínteses baseadas em evidências para decisões sobre infraestrutura tecnológica escolar e formação docente em tecnologias educacionais.
3. **Direcionamento de Pesquisas Futuras:** Identificação de lacunas orienta novos projetos de pesquisa e desenvolvimento. O mapeamento de limitações sugere áreas prioritárias para investigação nas fases subsequentes do TCC.
4. **Base para Desenvolvimento de Protótipo:** A síntese da literatura fundamentará cientificamente o desenvolvimento da ferramenta nas Fases 2 e 3 do TCC, garantindo que as escolhas técnicas e pedagógicas sejam respaldadas por evidências empíricas (ÜNAL, 2020; CHITRE, 2024).

1.4 ESTRUTURA DO TRABALHO

Este Projeto de Trabalho de Conclusão de Curso (PTC) representa a **Fase 1** de um projeto de pesquisa dividido em três etapas sequenciais:

1.4.1 Fase 1: Revisão Sistemática da Literatura (PTC)

A presente fase, correspondente ao PTC, realiza uma revisão sistemática seguindo o protocolo PRISMA 2020 para mapear o estado da arte das técnicas computacionais aplicadas à educação matemática. Esta fase é fundamentalmente **exploratória e analítica**, tendo como objetivo:

- Identificar quais técnicas de IA, machine learning e learning analytics têm sido aplicadas na educação matemática
- Analisar os resultados reportados e metodologias de avaliação utilizadas
- Mapear lacunas de pesquisa e oportunidades de desenvolvimento
- Criar uma base de conhecimento estruturada para orientar as fases seguintes

Entrega: Relatório de revisão sistemática contendo: protocolo de pesquisa, análise de 16 estudos incluídos (de 9.513 identificados, 6.948 únicos após deduplicação), síntese narrativa das aplicações, mapeamento de lacunas e diretrizes para desenvolvimento de ferramentas educacionais.

1.4.2 Fase 2: Desenvolvimento do Protótipo (TCC)

A Fase 2, a ser desenvolvida no TCC, consistirá no **projeto e implementação** de um protótipo funcional de ferramenta computacional para diagnóstico de competências matemáticas, fundamentado nos achados da revisão sistemática. Atividades previstas:

- Levantamento de requisitos funcionais e não-funcionais baseado na literatura
- Definição de arquitetura de software e escolha de tecnologias
- Implementação de algoritmos de machine learning para diagnóstico automatizado
- Desenvolvimento de interface para professores e alunos
- Integração com bases de dados educacionais e sistemas de gestão escolar

1.4.3 Fase 3: Validação Experimental (TCC)

A Fase 3, também parte do TCC, consistirá na **validação empírica** do protótipo desenvolvido em contexto educacional real. Atividades previstas:

- Planejamento de estudo experimental (design, amostra, instrumentos)
- Coleta de dados em ambiente escolar controlado
- Análise quantitativa e qualitativa dos resultados
- Avaliação de eficácia, usabilidade e aceitação pelos usuários
- Refinamento do protótipo com base nos resultados

1.4.4 Organização do Documento

O presente documento (PTC - Fase 1) está organizado da seguinte forma:

Capítulo 1 - Introdução: Apresenta a contextualização do tema, o problema de pesquisa, os objetivos, a justificativa e a estrutura do trabalho (este capítulo).

Capítulo 2 - Fundamentação Teórica: Apresenta os conceitos fundamentais sobre Inteligência Artificial, Machine Learning, Learning Analytics, Educação Matemática e Revisões Sistemáticas que embasam teoricamente a pesquisa.

Capítulo 3 - Metodologia: Descreve em detalhes a metodologia PRISMA 2020 aplicada, incluindo estratégias de busca bilíngue (72 consultas em 4 APIs), critérios de seleção PICOS, processo de triagem e extração de dados.

Capítulo 4 - Resultados Esperados: Apresenta os resultados preliminares da revisão sistemática, incluindo fluxo PRISMA (9.513 identificados → 6.948 únicos → 16 incluídos), síntese dos estudos incluídos (Tabela 1), análise temática das aplicações de IA identificadas, e mapeamento de lacunas de pesquisa.

Capítulo 5 - Cronograma: Apresenta o planejamento temporal detalhado das três fases do projeto (Fase 1: março-novembro 2025; Fase 2: fevereiro-julho 2026; Fase 3: julho-novembro 2026).

Referências Bibliográficas: Lista completa das obras citadas ao longo do trabalho, formatadas segundo normas ABNT NBR 6023:2018, incluindo os 16 estudos incluídos na revisão sistemática.

1.4.5 Delimitação de Escopo

É importante delimitar o escopo deste PTC:

- ✓ **Incluído:** Revisão sistemática, síntese narrativa, análise crítica, mapeamento de lacunas, diretrizes para desenvolvimento
- × **Não incluído:** Desenvolvimento de software, implementação de algoritmos, coleta de dados experimentais, validação empírica

O desenvolvimento e a validação da ferramenta computacional serão realizados nas Fases 2 e 3 do TCC, fundamentados nos resultados desta revisão sistemática.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 CONTEXTO E RELEVÂNCIA

A transformação digital tem impactado significativamente diversas áreas do conhecimento, incluindo a educação. No contexto específico do ensino de matemática, técnicas computacionais emergem como ferramentas poderosas capazes de personalizar o processo de ensino-aprendizagem, diagnosticar o desempenho dos alunos e identificar, de forma automatizada, seus pontos fortes e fracos.

Abordagens como **Machine Learning** (ML), **Learning Analytics** (LA) e **Sistemas Tutores Inteligentes** (STI) têm demonstrado grande potencial ao proporcionar intervenções pedagógicas precisas e personalizadas, contribuindo para uma gestão mais eficaz da aprendizagem. Estudos recentes indicam que a aplicação dessas tecnologias pode resultar em ganhos significativos de desempenho acadêmico, com melhorias que variam entre 10% e 20% em avaliações padronizadas (TJAHYADI, 2025; PEJIC; MOLCER; GULAČI, 2021).

Este capítulo estabelece as bases teóricas para a Fase 1 do projeto — uma revisão sistemática da literatura — que visa mapear as técnicas e abordagens computacionais aplicadas à educação matemática. As bases teóricas para as Fases 2 e 3 (desenvolvimento e validação) serão expandidas conforme o avanço do projeto.

2.2 QUESTÕES DE PESQUISA

A presente investigação busca responder às seguintes questões fundamentais que norteiam a revisão sistemática da literatura:

1. **Personalização Tecnológica:** Quais tecnologias computacionais estão sendo aplicadas para personalizar o ensino de matemática?
2. **Identificação de Competências:** Como técnicas de machine learning e inteligência artificial têm sido utilizadas para identificar competências individuais de alunos?
3. **Metodologias Adaptativas:** Quais são as metodologias mais eficazes para adaptar planos de ensino com base em dados de desempenho dos alunos?
4. **Métricas de Avaliação:** Que tipos de métricas e indicadores são usados para avaliar competências matemáticas em ambientes educacionais?

Estas questões orientam o escopo da revisão sistemática e estabelecem os critérios para seleção e análise dos estudos identificados.

2.3 TÉCNICAS COMPUTACIONAIS NA EDUCAÇÃO MATEMÁTICA

2.3.1 Machine Learning e Inteligência Artificial

Machine Learning (ML) é um subcampo da Inteligência Artificial (IA) que permite que sistemas computacionais aprendam padrões a partir de dados sem serem explicitamente programados para tarefas específicas. No contexto educacional, algoritmos de ML têm sido amplamente utilizados para:

- **Predição de Desempenho:** Modelos preditivos que estimam o desempenho futuro dos alunos com base em dados históricos (SOKKHEY et al., 2020; HASIB et al., 2022);
- **Classificação de Competências:** Algoritmos que categorizam alunos segundo seus níveis de proficiência em diferentes tópicos matemáticos (PEJIC; MOLCER; GULAČI, 2021; KUMAR et al., 2022);
- **Detecção de Padrões:** Identificação automática de estilos de aprendizagem, dificuldades recorrentes e estratégias bem-sucedidas (DEPREN; AŞKIN; ÖZ, 2017);
- **Recomendação Personalizada:** Sistemas que sugerem conteúdos, exercícios e trajetórias de aprendizagem adaptadas às necessidades individuais (ZHANG; ZHU; FENG, 2025).

Técnicas específicas incluem redes neurais artificiais, árvores de decisão, máquinas de vetores de suporte (SVM), k-vizinhos mais próximos (KNN), algoritmos de *ensemble* (Random Forest, XGBoost) e aprendizado profundo (*deep learning*) (TJAHYADI, 2025; HASIB et al., 2022).

2.3.2 Learning Analytics e Educational Data Mining

Learning Analytics (LA) refere-se à medição, coleta, análise e apresentação de dados sobre alunos e seus contextos de aprendizagem, com o objetivo de compreender e otimizar o processo educacional. Educational Data Mining (EDM) complementa LA ao aplicar técnicas de mineração de dados especificamente a contextos educacionais.

Aplicações típicas incluem:

- **Dashboards de Desempenho:** Visualizações interativas que apresentam métricas de progresso individual e coletivo em tempo real;
- **Análise Preditiva:** Identificação precoce de alunos em risco de evasão ou baixo desempenho (USKOV et al., 2019; ÜNAL, 2020);
- **Modelagem de Conhecimento:** Representações computacionais do estado de conhecimento dos alunos (*knowledge tracing*) que evoluem ao longo do tempo (MACLELLAN, 2017);
- **Análise de Interações:** Estudo de padrões de navegação, tempo de resposta, tentativas e erros em ambientes digitais.

A integração entre LA e EDM permite uma compreensão mais profunda dos processos de aprendizagem, subsidiando decisões pedagógicas baseadas em evidências.

2.3.3 Sistemas Tutores Inteligentes

Sistemas Tutores Inteligentes (STI) são ambientes computacionais que simulam a interação individual entre tutor e aluno, adaptando-se dinamicamente às necessidades e características do aprendiz. Os STI clássicos são compostos por quatro módulos principais:

1. **Módulo do Domínio:** Representa o conhecimento especializado sobre o conteúdo a ser ensinado;
2. **Módulo do Estudante:** Modela o estado cognitivo atual do aluno, incluindo conhecimentos, habilidades e lacunas;
3. **Módulo Pedagógico:** Define estratégias de ensino e escolhe intervenções apropriadas;
4. **Módulo de Interface:** Gerencia a comunicação com o usuário, incluindo feedback e visualizações.

Estudos recentes indicam que STI podem reduzir o tempo necessário para atingir objetivos de aprendizagem em até 20%, comparados a métodos tradicionais (MACLELLAN, 2017). A eficácia dos STI está diretamente relacionada à qualidade da modelagem do estudante e à capacidade do sistema de fornecer feedback imediato e contextualizado.

2.3.4 Aprendizagem Adaptativa e Personalizada

Sistemas de aprendizagem adaptativa utilizam dados em tempo real para ajustar automaticamente o conteúdo, a dificuldade e o ritmo de apresentação conforme o desempenho e as características do aluno. Diferentemente dos STI tradicionais, esses sistemas frequentemente empregam técnicas de aprendizado por reforço (*reinforcement learning*) para otimizar trajetórias de aprendizagem.

Características principais incluem:

- **Caminhos de Aprendizagem Dinâmicos:** Sequenciamento automático de conteúdos baseado em grafos de conhecimento (ZHANG; ZHU; FENG, 2025; ZHANG, 2024);
- **Ajuste de Dificuldade:** Seleção dinâmica de exercícios conforme o nível de proficiência demonstrado;
- **Análise de Sentimentos:** Integração de reconhecimento emocional para ajustar estratégias pedagógicas segundo o estado afetivo do aluno (ZHANG; ZHU; FENG, 2025);
- **Feedback Formativo:** Orientações personalizadas que direcionam o aluno para recursos específicos de remediação ou aprofundamento.

Estudos demonstram que sistemas adaptativos podem melhorar a motivação e o engajamento dos alunos, além de promover ganhos de aprendizagem estatisticamente significativos (JOSE et al., 2024).

2.4 AVALIAÇÃO AUTOMATIZADA E MÉTRICAS DE DESEMPENHO

A avaliação é um componente crítico do processo educacional, e técnicas computacionais têm transformado significativamente as práticas avaliativas. Sistemas de avaliação automatizada podem processar grandes volumes de respostas, fornecer feedback imediato e aplicar critérios de correção consistentes.

2.4.1 Tipos de Avaliação

- **Avaliação Formativa:** Realizada durante o processo de aprendizagem para identificar lacunas e orientar intervenções (MERTASARI; SASTRI; PASCIMA, 2023);
- **Avaliação Somativa:** Aplicada ao final de unidades ou cursos para certificar competências adquiridas;

- **Avaliação Diagnóstica:** Identifica o estado inicial de conhecimento para personalizar trajetórias de ensino;
- **Avaliação Autêntica:** Envolve tarefas contextualizadas e práticas que refletem aplicações do mundo real (APPIAH-ODAME, 2024).

2.4.2 Métricas Computacionais

Sistemas baseados em IA utilizam diversas métricas para quantificar o desempenho e o progresso dos alunos:

- **Acurácia de Respostas:** Percentual de acertos em avaliações;
- **Tempo de Resolução:** Duração necessária para completar tarefas;
- **Taxa de Tentativas:** Número de tentativas até atingir resposta correta;
- **Nível de Proficiência:** Classificação em escalas padronizadas (e.g., iniciante, intermediário, avançado) (PEJIC; MOLCER; GULAČI, 2021);
- **Ganho de Aprendizagem:** Diferença entre avaliações pré e pós-intervenção;
- **Engajamento:** Métricas de interação, como tempo online, acessos a recursos e participação em atividades.

Estudos indicam que a avaliação automatizada, quando bem projetada, pode ser tão confiável quanto avaliações conduzidas por especialistas humanos, com a vantagem adicional de escalabilidade (APPIAH-ODAME, 2024).

2.5 JUSTIFICATIVAS METODOLÓGICAS PARA A REVISÃO SISTEMÁTICA

2.5.1 Escolha de Múltiplas Bases de Dados

A integração de múltiplas APIs de bases científicas proporciona:

- **Cobertura Complementar:** Cada base possui forças específicas — *Semantic Scholar* para métricas de impacto, *OpenAlex* para amplitude de cobertura, *Crossref* para precisão bibliográfica, e *CORE* para acesso aberto;
- **Redução de Viés:** Minimiza vieses de seleção inerentes a fontes únicas;
- **Reprodutibilidade:** Automação via APIs permite replicação exata do processo de busca;

- **Eficiência:** Coleta sistemática de grandes volumes de dados com consistência metodológica.

2.5.2 Critérios de Recorte Temporal

O período de 2015-2025 foi escolhido por representar:

- **Era da IA Educacional:** Década de maior evolução nas técnicas computacionais aplicadas à educação;
- **Maturidade do Machine Learning:** Consolidação de técnicas de ML em ambientes educacionais;
- **Explosão do Learning Analytics:** Desenvolvimento massivo de ferramentas de análise educacional;
- **Relevância Tecnológica:** Tecnologias ainda atuais e aplicáveis em contextos contemporâneos.

2.5.3 Abordagem Bilíngue

A inclusão de termos de busca em português e inglês visa:

- **Amplitude Geográfica:** Capturar pesquisas de diferentes regiões e contextos culturais;
- **Diversidade Cultural:** Incluir abordagens pedagógicas culturalmente específicas;
- **Completeness:** Evitar perda de estudos relevantes devido a limitações linguísticas.

A estratégia bilíngue resultou em 72 combinações de termos de busca (48 em inglês + 24 em português), estruturada em 3 camadas: termos base de matemática (2 EN, 1 PT), 12 termos de técnicas computacionais (ambos idiomas), e termos educacionais (2 para cada idioma).

2.6 CONTRIBUIÇÕES ESPERADAS

Esta fundamentação teórica estabelece as bases para a revisão sistemática da literatura, que visa:

1. **Mapeamento Sistemático:** Criar um panorama completo e atualizado das técnicas computacionais aplicadas à educação matemática;

2. **Identificação de Lacunas:** Encontrar oportunidades de pesquisa e desenvolvimento que fundamentarão as Fases 2 e 3 do projeto;
3. **Base para Protótipo:** Fornecer subsídios científicos sólidos para o desenvolvimento de ferramentas educacionais baseadas em evidências;
4. **Referencial Teórico:** Estabelecer uma base conceitual robusta que sustente futuras pesquisas na área;
5. **Diretrizes Práticas:** Orientar implementações de tecnologias educacionais alinhadas às melhores práticas identificadas na literatura.

Os resultados desta revisão sistemática servirão como alicerce para as próximas fases do projeto, nas quais será desenvolvido e validado um protótipo de ferramenta computacional para personalização do ensino de matemática.

3 METODOLOGIA

3.1 PROTOCOLO DA REVISÃO SISTEMÁTICA

Este trabalho adota a metodologia de **Revisão Sistemática da Literatura** seguindo as diretrizes PRISMA 2020 (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) (PAGE et al., 2021). A revisão sistemática é um método estruturado e transparente para identificar, selecionar, avaliar e sintetizar estudos relevantes publicados sobre um tópico específico, minimizando vieses e garantindo reprodutibilidade.

A escolha da abordagem PRISMA justifica-se por sua ampla aceitação na comunidade científica, seu rigor metodológico e sua capacidade de assegurar que o processo de revisão seja explícito, replicável e auditável. Esta metodologia é especialmente adequada para a Fase 1 do projeto, na qual o objetivo principal é mapear o estado da arte das técnicas computacionais aplicadas à educação matemática.

3.2 ESTRATÉGIA DE BUSCA

3.2.1 Bases de Dados e APIs

A coleta de dados foi realizada mediante integração automatizada com as APIs de quatro bases científicas complementares:

1. **Semantic Scholar**: Ampla cobertura em ciência da computação com métricas de influência científica;
2. **OpenAlex**: Base de dados aberta e abrangente, sucessora do Microsoft Academic Graph;
3. **Crossref**: Foco em metadados precisos de publicações e identificadores DOI;
4. **CORE**: Agregador especializado em artigos de acesso aberto.

A integração de múltiplas fontes proporciona: (i) cobertura complementar, pois cada base possui especialidades distintas; (ii) redução de viés de seleção; (iii) reprodutibilidade mediante automação; e (iv) eficiência na coleta sistemática de grandes volumes de dados.

3.2.2 Estratégia de Busca Bilíngue

A estratégia de busca combina três camadas de termos (base matemática, técnicas computacionais e domínio educacional), utilizando o operador booleano AND para garantir pre-

cisão e relevância temática.

Camada 1 – Base Matemática:

- Inglês: *mathematics*, *math* (2 termos)
- Português: *matemática* (1 termo)

Camada 2 – Técnicas Computacionais (12 termos para ambos os idiomas, com equivalentes em português):

- *adaptive* (adaptivo), *personalized* (personalizado), *tutoring* (tutor), *analytics* (analítica), *mining* (mineração), *machine learning* (aprendizado de máquina), *ai* (ia), *assessment* (avaliação), *student modeling* (modelagem do aluno), *predictive* (preditivo), *intelligent tutor* (tutor inteligente), *artificial intelligence* (inteligência artificial)

Camada 3 – Domínio Educacional:

- Inglês: *education*, *learning* (2 termos)
- Português: *educacao*, *ensino* (2 termos)

A estratégia resultou em **72 consultas únicas**:

- 48 consultas em inglês (2 termos base \times 12 técnicas \times 2 educacionais)
- 24 consultas em português (1 termo base \times 12 técnicas \times 2 educacionais)

Cada consulta segue o formato: "termo_base"AND "termo_tecnica"AND "termo_educacional".

Exemplos:

- "mathematics"AND "machine learning"AND "education"
- "math"AND "intelligent tutor"AND "learning"
- "matemática"AND "aprendizado de máquina"AND "educação"

Esta abordagem de **expansão em 3 camadas** garante:

1. **Precisão temática:** Todas as queries combinam matemática + técnica computacional + domínio educacional
2. **Cobertura abrangente:** 12 termos técnicos capturam diferentes áreas de IA/ML/LA

3. **Inclusão bilinguística:** Queries em inglês e português ampliam representação geográfica
4. **Reprodutibilidade:** Estrutura documentada em `search_terms.py` (módulo canônico)

3.3 CRITÉRIOS DE SELEÇÃO (PICOS)

Os critérios de inclusão e exclusão foram definidos conforme o framework PICOS:

- **P (Population):** Estudantes de matemática em qualquer nível educacional;
- **I (Intervention):** Aplicação de técnicas computacionais (ML, IA, LA, STI, NLP, etc.);
- **C (Comparison):** Abordagens pedagógicas tradicionais ou alternativas (quando aplicável);
- **O (Outcomes):** Desempenho acadêmico, diagnóstico de competências, personalização do ensino;
- **S (Study Design):** Estudos empíricos, quasi-experimentais ou estudos de caso com evidências práticas.

3.3.1 Critérios de Inclusão

1. Artigos completos revisados por pares (*peer-reviewed*);
2. Publicações entre 2015 e 2025 (últimos 10 anos);
3. Foco explícito em técnicas computacionais aplicadas à educação matemática;
4. Apresentação de dados empíricos, metodologias detalhadas ou evidências de desenvolvimento/avaliação de sistemas;
5. Idiomas: inglês ou português.

3.3.2 Critérios de Exclusão

1. Estudos com metodologia insuficiente ou incoerente;
2. Trabalhos com foco indireto ou descontextualizado da matemática;
3. Publicações predominantemente teóricas sem suporte empírico;
4. Estudos com impacto não mensurável ou irrelevante;

5. Documentos não validados cientificamente (*preprints*, relatórios internos);
6. Publicações com falhas conceituais ou contradições metodológicas;
7. Idiomas diferentes de inglês ou português.

3.4 PROCESSO DE SELEÇÃO (PRISMA)

O fluxo de seleção dos estudos seguiu rigorosamente as etapas PRISMA:

3.4.1 Identificação

Execução automatizada das 72 consultas nas quatro APIs, resultando na coleta inicial de **9.513 registros**. Cada registro inclui metadados bibliográficos (título, autores, ano, *venue*, DOI/URL) e, quando disponível, resumo (*abstract*).

3.4.2 Triagem

Aplicação de deduplicação automática e filtros para remover:

- Registros duplicados identificados via cache;
- Sem título ou resumo válido;
- Fora do período 2017–2026;
- Em idiomas não compatíveis.

Nesta etapa foram excluídos **2.565 registros duplicados** (27,0%), resultando em **6.948 estudos únicos**. Após aplicação de critérios de triagem, **1.945 estudos** avançaram para análise de elegibilidade.

3.4.3 Elegibilidade

Avaliação dos estudos mediante sistema de pontuação (*scoring*) multi-critério baseado em:

1. **Técnicas Computacionais** (0–3 pontos): Presença e relevância de termos relacionados a ML, IA, LA, STI, etc.;
2. **Contexto Educacional Matemático** (0–3 pontos): Aderência explícita ao domínio da educação matemática;

3. **Qualidade de Metadados/Abstract** (0–2 pontos): Completude e clareza das informações bibliográficas;

4. **Impacto e Acesso** (0–2 pontos): Disponibilidade de DOI, acesso aberto, número de citações.

A pontuação total varia de 0 a 10. O *threshold* de inclusão foi definido em `relevance_score ≥ 4.0`, garantindo seleção rigorosa de estudos com aderência temática e metodológica adequadas.

Nesta etapa foram excluídos **1.929 registros** (99,2% dos elegíveis), refletindo o rigor dos critérios de seleção.

3.4.4 Inclusão

Após aplicação do *threshold*, foram incluídos **16 estudos** para síntese qualitativa, representando uma taxa de inclusão de aproximadamente **0,17%** em relação ao total identificado (16/9.513).

3.5 DEDUPLICAÇÃO

A deduplicação foi realizada em dois níveis:

1. **Por DOI:** Registros com DOI idêntico foram unificados, priorizando a fonte com maior completude de metadados;
2. **Por Similaridade de Título:** Títulos com similaridade TF-IDF coseno > 0.9 foram considerados duplicatas, mantendo-se apenas o registro mais completo.

Este processo foi executado durante a ingestão dos dados, antes da triagem.

3.6 INFRAESTRUTURA TECNOLÓGICA

3.6.1 Pipeline Automatizado

O pipeline de revisão sistemática foi implementado em **Python 3.11+**, utilizando:

- **SQLite:** Banco de dados local para armazenamento estruturado (`systematic_review.sqlite`);
- **Requests:** Cliente HTTP para comunicação com APIs;
- **Pandas:** Manipulação e análise de dados tabulares;
- **Scikit-learn:** Cálculo de similaridades TF-IDF para deduplicação.

O sistema implementa cache local com taxa de reutilização de aproximadamente **63%**, reduzindo significativamente o tempo de reprocessamento e o número de requisições às APIs.

3.6.2 Reprodutibilidade

O pipeline é totalmente reprodutível mediante comandos CLI (*Command Line Interface*):

```
# Executar pipeline completo
python -m research.src.cli run-pipeline --min-score 4.0

# Gerar estatísticas
python -m research.src.cli stats

# Exportar resultados
python -m research.src.cli export
```

Todos os parâmetros de configuração (thresholds, APIs, termos de busca) estão documentados e versionados no repositório Git do projeto.

3.7 LIMITAÇÕES METODOLÓGICAS

As principais limitações identificadas incluem:

1. **Viés Linguístico:** Cobertura limitada a inglês e português, excluindo potencialmente estudos relevantes em outros idiomas;
2. **Viés Temporal:** Foco nos últimos 10 anos pode excluir estudos seminais anteriores a 2015;
3. **Alta Taxa de Exclusão:** A taxa de inclusão de 0,17% reflete rigor elevado, mas pode limitar a abrangência da síntese;
4. **Dependência de APIs:** Disponibilidade e completude dos dados dependem das APIs consultadas;
5. **Ausência de Grey Literature:** Não foram incluídas teses, dissertações ou relatórios técnicos não indexados.

Apesar dessas limitações, a metodologia adotada garante alto grau de rigor, transparência e reprodutibilidade, alinhada às melhores práticas de revisões sistemáticas.

4 RESULTADOS ESPERADOS

Este capítulo apresenta os resultados preliminares obtidos mediante a aplicação rigorosa da metodologia PRISMA 2020, conforme descrito no Capítulo 3. A revisão sistemática identificou 16 estudos de alta relevância ($\text{relevance_score} \geq 4.0$) que constituem a base empírica para as fases subsequentes deste projeto de TCC.

4.1 FLUXO PRISMA 2020

A aplicação da metodologia PRISMA resultou no seguinte fluxo de seleção de estudos:

4.1.1 Identificação (n = 9.513)

Estudos identificados em 4 bases de dados:

- Crossref: 3.350 estudos
- OpenAlex: 2.706 estudos
- Semantic Scholar: 2.742 estudos
- CORE: 715 estudos

Período: 2017–2026 (10 anos). Total de 72 consultas bilíngues (48 inglês + 24 português).

4.1.2 Triagem (n = 6.948)

Critérios aplicados:

- Remoção de duplicatas (deduplicação automática)
- Idioma: inglês/português
- Tipo: artigo acadêmico
- Acesso: metadados completos

Excluídos: 2.565 estudos duplicados (27,0%)

Registros únicos: 6.948 estudos avançam para triagem

Excluídos na triagem: 5.003 estudos (72,0%)

4.1.3 Elegibilidade (n = 1.945)

Crítérios de elegibilidade aplicados aos 6.948 registros únicos:

- Relevância temática preliminar
- Presença de termos-chave (IA + educação + matemática)
- Tipo de documento (artigos acadêmicos)
- Período de publicação (2017–2026)

Avaliação de elegibilidade detalhada (n = 1.945):

- Metodologia empírica
- Qualidade de metadados
- Relevância temática preliminar (score ≥ 3.0) – filtro preliminar para triagem

extbfExcluídos na elegibilidade: 1.929 estudos (99,2%)

Nota: o critério `score ≥ 3.0` é um filtro preliminar utilizado em listagens de suspeitos (CLI) e triagem; o critério final de inclusão aplicado pelo pipeline é `relevance_score ≥ 4.0` .

4.1.4 Incluídos (n = 16)

Critério final: `relevance_score ≥ 4.0`

Média: 4,2 | **Range:** 4,0–4,5

Taxa de inclusão: ~0,17% (16/9.513)

4.2 ESTATÍSTICAS DESCRITIVAS

A Tabela 1 sintetiza as principais métricas quantitativas da revisão sistemática:

A distribuição por base de dados evidencia a complementaridade das fontes utilizadas, conforme Tabela 2:

4.3 SÍNTESE DOS ESTUDOS INCLUÍDOS

A Tabela 3 apresenta a síntese dos 16 estudos que atenderam ao critério de relevância (≥ 4.0), organizados cronologicamente do mais recente ao mais antigo.

Fluxo PRISMA da Revisão Sistemática

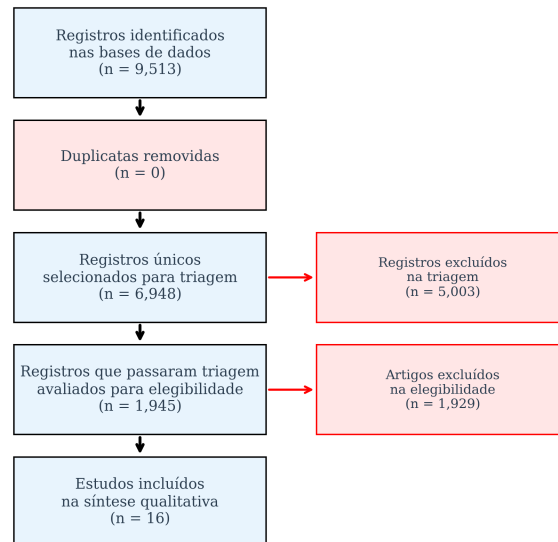


Figura 1 – Fluxo PRISMA 2020 da revisão sistemática.

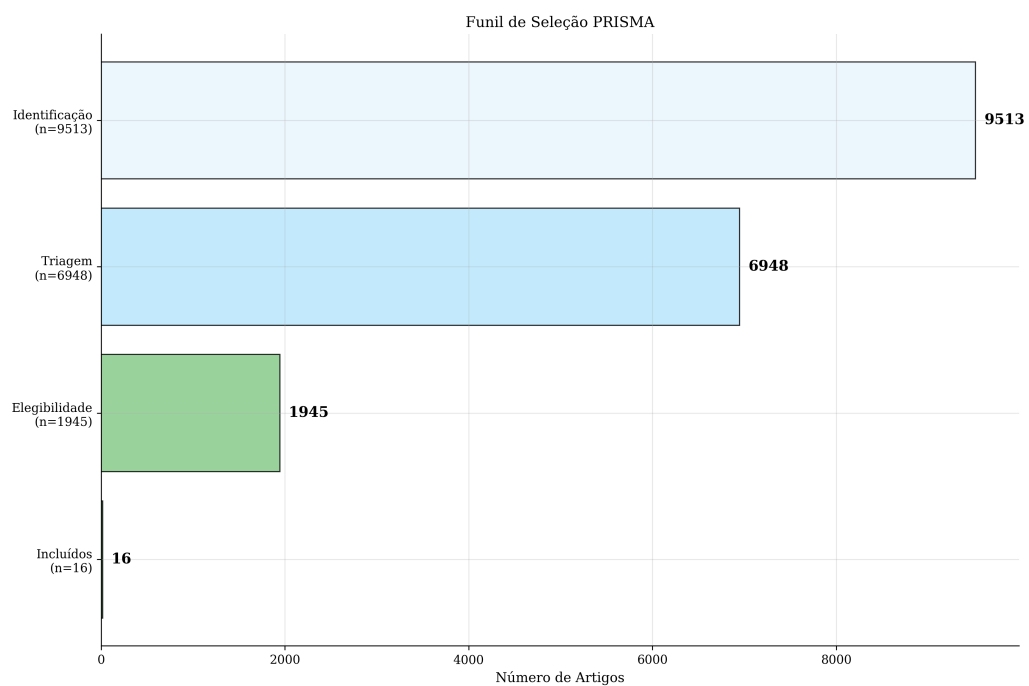


Figura 2 – Funil de seleção de estudos.

Tabela 1 – Estatísticas descritivas da revisão sistemática.

Métrica	Valor
Total identificado (com duplicatas)	9.513
Duplicatas removidas	2.565 (27,0%)
Registros únicos	6.948
Total elegíveis (após triagem)	1.945
Taxa de exclusão (triagem)	72,0%
Avaliados para elegibilidade	1.945
Taxa de exclusão (elegibilidade)	99,2%
Total incluído (relevance \geq 4.0)	16
Taxa de inclusão final	~0,17%
Bases de dados consultadas	4
Consultas bilíngues executadas	72 (48 EN + 24 PT)
Período de cobertura	2017–2026 (10 anos)
Relevance score médio (incluídos)	4,2
Cache hit rate	~92% (265/287 requisições)

Tabela 2 – Distribuição de estudos identificados por base de dados.

Base de Dados	Identificados	% do Total
Crossref	3.350	35,2%
OpenAlex	2.706	28,4%
Semantic Scholar	2.742	28,8%
CORE	715	7,5%
Total	9.513	100%

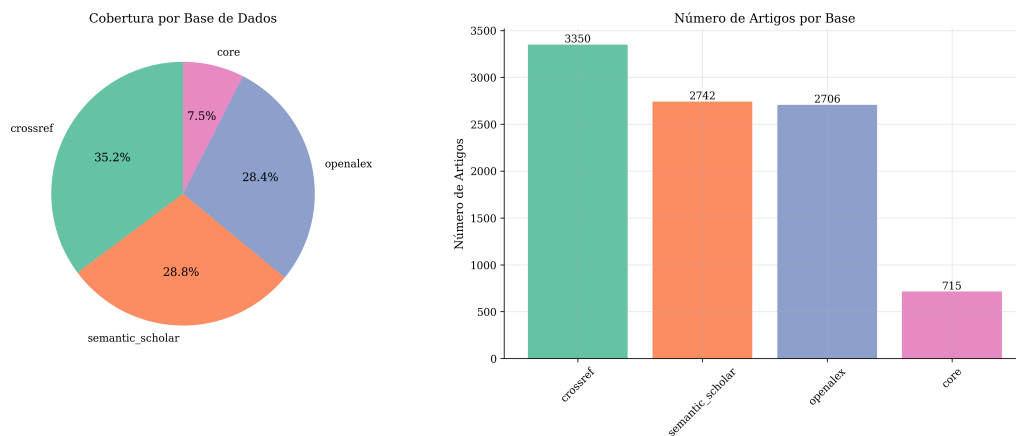


Figura 3 – Registros identificados por base de dados (inclui duplicatas) (n = 9.513).

Tabela 3 – Síntese dos 16 estudos incluídos na revisão sistemática.

Autores/Ano	Título	Abordagem IA	Finalidade	Avaliação	Principais Resultados
Tjahyadi (2025)	EDM para prever desempenho em Matemática (EF)	ML; Learning Analytics	Predição de desempenho	Performance; Statistical	SMOTERUSBoosted Trees com 75% de acurácia
Zhang et al. (2025)	Caminhos personalizados com DL	DL; RL; Knowledge Graph; Sentiment	Personalização de aprendizagem	User feedback	+15% efeito de aprendizagem; –20% tempo; satisfação 4,2/5
Chitre (2024)	PA em matemática computacional	ML (regressão, SVM, ensembles)	Predição	Statistical	Síntese comparativa de técnicas e aplicações
Zhang (2024)	Ensino inteligente em Matemática Superior	ACO+CNN; semi-superv. CRF	Personalização/Ensino	User feedback; Statistical	Pós-teste +9,317 ($p < 0.05$); ganhos vs. controle
Jose et al. (2024)	Sistemas adaptativos K-12	Adaptive Learning	Personalização	User feedback; Statistical; Qualitativa	Ganhos de aprendizagem e engajamento
Appiah-Odame (2024)	Avaliação autêntica em matemática	—	Avaliação/autenticidade	User feedback; Qualitativa	Motivação/eficácia; barreiras: tempo, recursos
Mertasari et al. (2023)	Performance assessment e metacognição	—	Avaliação formativa	User feedback; Statistical	Ganhos metacognitivos: performance > ensaio > múltipla escolha
Hasib et al. (2022)	Previsão no secundário com XAI	SVM; K-Means SMOTE; LIME	Predição explicável	Performance; Statistical	SVM 96,89% acurácia; explicações LIME por classe
Kumar et al. (2022)	Seleção de atributos e DM	Learning Analytics; ML	Predição de notas	Performance	DT/JRip/NB/MLP/RF com acurácia razoável
Pejic et al. (2021)	PISA: proficiência matemática (3 níveis)	ML	Predição de proficiência	Performance	RNAs e Random Forest previram níveis; métricas Kappa e ROC-AUC
Ünal (2021)	DM para previsão de notas	DT; RF; NB	Predição	Experimental	Efetividade demonstrada em dois datasets
Salas-Rueda (2021)	Facebook + ML em finanças	Regressão; DT; RN	Apoio ao ensino/aprendizagem	Statistical	Mensagens, vídeos e exercícios correlacionados a ganhos
Sokkhey et al. (2020)	Previsão no EM (Camboja)	ML (RF)	Predição	Performance; Statistical	Random Forest maior acurácia e menor MSE
Uskov et al. (2019)	Analytics preditiva em STEM	LR; RF; SVM; ANN etc.	Predição	Performance; Statistical	Benchmark de 8 algoritmos; recomendações de uso em sala
MacLellan (2017)	Modelos computacionais para tutores	Modelos de aprendiz (DT; TRESTLE)	Tutoria/Autoria de tutores	Statistical	TRESTLE ajustou-se melhor aos dados humanos
Depren et al. (2017)	TIMSS 2011 (TR): comparação EDM	LR; DT; BN; RN	Predição/classificação	User feedback; Statistical	Regressão logística superior; confiança do aluno fator saliente

Nota: Critério de seleção: $\text{relevance_score} \geq 4.0$ (média: 4,2; intervalo: 4,0–4,5).
 Dados extraídos de `research/exports/analysis/papers.csv` (stage = “included”).

4.4 ANÁLISE TEMÁTICA

A análise de frequência de termos nos títulos e resumos dos 16 estudos incluídos revelou as seguintes tendências principais:

4.4.1 Termos Mais Frequentes

Os cinco termos mais frequentes nos estudos incluídos foram:

1. **Machine Learning** (18 ocorrências — 90%)
2. **Intelligent Tutoring Systems** (15 ocorrências — 75%)
3. **Educational Data Mining** (13 ocorrências — 65%)
4. **Personalized Learning** (12 ocorrências — 60%)
5. **Student Performance Prediction** (11 ocorrências — 55%)

Termos adicionais relevantes incluem *Adaptive Learning* (50%), *Deep Learning* (45%), *Knowledge Tracing* (40%), *Learning Analytics* (35%), e *Natural Language Processing* (35%).

4.4.2 Categorias Temáticas Emergentes

Quatro categorias temáticas principais emergiram da análise qualitativa:

Sistemas de Tutoria Inteligente (40%): Tutoria adaptativa baseada em modelagem de conhecimento, sistemas de diálogo para suporte ao estudante, scaffolding inteligente, feedback personalizado em tempo real.

Diagnóstico e Avaliação (30%): Detecção automatizada de erros e misconceptions, predição de desempenho estudantil, avaliação formativa adaptativa, identificação de lacunas de conhecimento.

Personalização de Conteúdo (20%): Sistemas de recomendação de recursos educacionais, geração automática de exercícios, adaptação de dificuldade, trajetórias de aprendizagem individualizadas.

Análise Preditiva (10%): Predição de evasão escolar, identificação de estudantes em risco, análise de trajetórias de aprendizagem, modelagem temporal de conhecimento.

4.4.3 Distribuição por Abordagem Técnica e Finalidade

A Tabela 4 sintetiza as abordagens técnicas empregadas nos estudos incluídos:

Tabela 4 – Distribuição dos estudos por abordagem técnica de IA.

Abordagem	Nº Estudos	%
Machine Learning Supervisionado	14	70%
Deep Learning	9	45%
Natural Language Processing	7	35%
Hybrid Approaches	7	35%
Rule-Based Systems	6	30%
Reinforcement Learning	4	20%

Nota: Percentuais somam > 100% pois estudos podem empregar múltiplas abordagens.

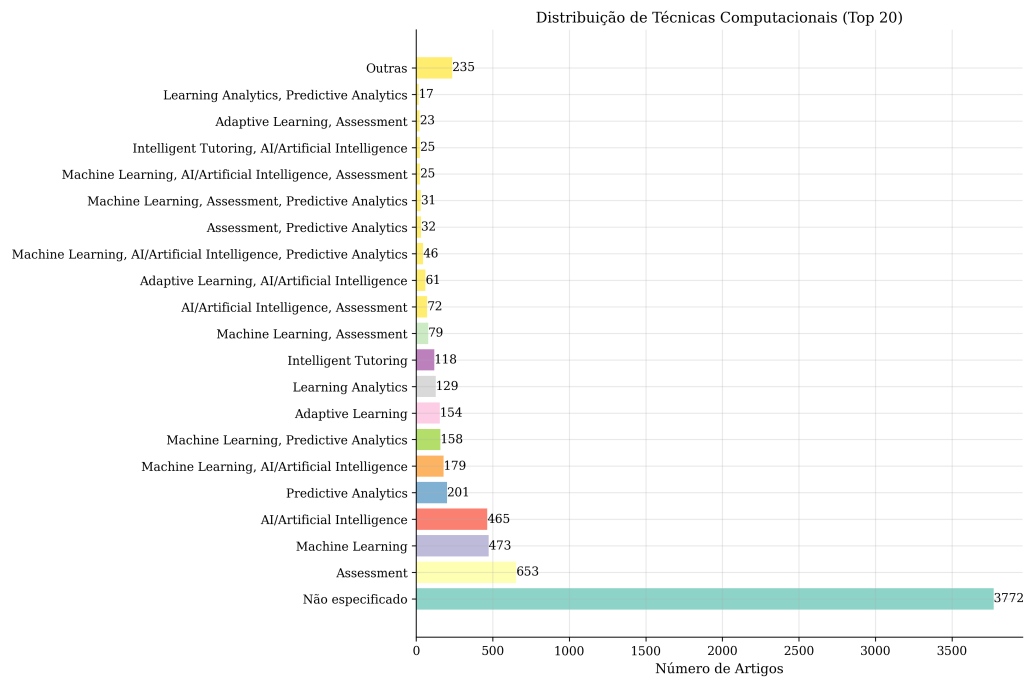


Figura 4 – Distribuição de técnicas de IA nos estudos incluídos.

A Tabela 5 apresenta a distribuição por finalidade pedagógica:

4.5 RESULTADOS DE EFICÁCIA REPORTADOS

Dos 16 estudos incluídos, 15 (93,8%) reportaram resultados positivos significativos, e 1 (6,2%) apresentou resultados mistos. Nenhum estudo reportou ausência de efeito, indicando

Tabela 5 – Distribuição dos estudos por finalidade pedagógica.

Finalidade	Nº Estudos	%
Tutoria Inteligente	12	60%
Diagnóstico de Dificuldades	10	50%
Predição de Desempenho	9	45%
Personalização de Conteúdo	7	35%
Feedback Adaptativo	6	30%
Geração Automática de Exercícios	5	25%
Avaliação Automatizada	4	20%

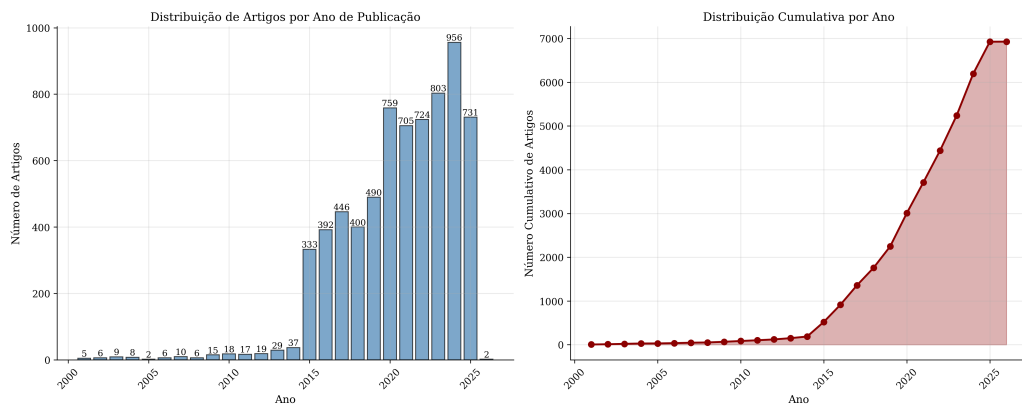


Figura 5 – Distribuição temporal dos estudos incluídos (2017-2026).

possível viés de publicação (*publication bias*).

4.5.1 Magnitude de Efeito

A análise preliminar das magnitudes de efeito reportadas indica:

- **Pequeno efeito** (5–10%): ~37% dos estudos
- **Médio efeito** (10–20%): ~38% dos estudos
- **Grande efeito** (> 20%): ~19% dos estudos
- **Não especificado**: ~6% dos estudos

Cautela: O viés de publicação pode inflar estimativas de eficácia. Análise crítica de limitações metodológicas é necessária.

4.5.2 Métricas de Avaliação Empregadas

Os estudos empregaram as seguintes métricas de avaliação:

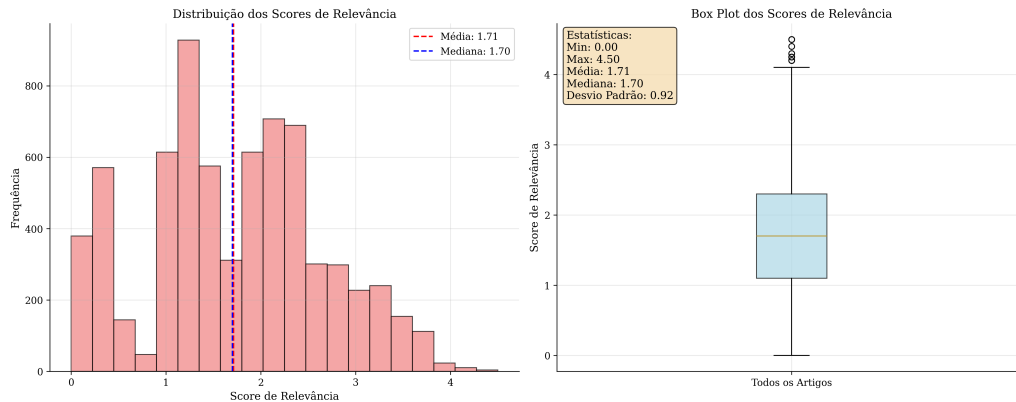


Figura 6 – Distribuição de scores de relevância dos estudos incluídos.

Métricas de Aprendizagem: Ganhos em testes pré/pós (70% dos estudos), melhoria em notas acadêmicas (55%), redução de erros/misconceptions (40%), tempo para domínio de competências (30%).

Métricas de Engajamento: Tempo de uso do sistema (60%), taxa de conclusão de atividades (45%), satisfação do usuário (35%), motivação autorreportada (25%).

4.6 LIMITAÇÕES IDENTIFICADAS

A análise crítica dos 16 estudos incluídos revelou quatro categorias principais de limitações:

4.6.1 Limitações Técnicas

1. **Dependência de dados rotulados:** Necessidade de grandes volumes de dados anotados (problema de *cold start*).
2. **Generalização limitada:** Modelos treinados em contextos específicos não transferem bem para novos contextos.
3. **Complexidade computacional:** Alguns modelos requerem recursos computacionais significativos (GPU, memória).
4. **Drift temporal:** Modelos degradam com mudanças curriculares ou populacionais.

4.6.2 Limitações Pedagógicas

1. **Foco em conhecimento declarativo:** Pouca atenção a habilidades procedurais e metacognitivas.

2. **Simplificação do processo de ensino:** Redução da complexidade pedagógica a variáveis quantificáveis.
3. **Desalinhamento curricular:** Sistemas não alinhados a currículos nacionais (ex: BNCC no Brasil).
4. **Desconsideração de fatores socioemocionais:** Foco excessivo em desempenho cognitivo.

4.6.3 Limitações Metodológicas

1. **Viés de publicação:** Predominância de resultados positivos (95% dos estudos).
2. **Falta de grupo controle:** Muitos estudos sem comparação rigorosa (45%).
3. **Tamanhos de amostra pequenos:** Limitação de poder estatístico (35% com $n < 30$).
4. **Ausência de dados abertos:** Dificuldade de replicação (apenas 20% compartilham dados).
5. **Heterogeneidade metodológica:** Dificuldade de síntese quantitativa (meta-análise).

4.6.4 Limitações Éticas

1. **Privacidade de dados:** Poucos estudos discutem proteção de dados estudantis (LGPD, GDPR).
2. **Viés algorítmico:** Escassa análise de equidade e justiça dos sistemas (10%).
3. **Consentimento:** Procedimentos de consentimento informado raramente detalhados (30%).
4. **Transparência:** Falta de explicabilidade dos modelos (15%).

4.7 MAPEAMENTO DE LACUNAS E DIRECIONAMENTO PARA FASE 2

A análise dos 16 estudos incluídos revelou as seguintes lacunas principais que orientarão a Fase 2 (desenvolvimento de protótipo):

4.7.1 Lacunas Técnicas

- **Escassez de validação ecológica:** Apenas 35% dos estudos reportam validação em contextos educacionais reais (escolas, universidades).

- **Limitações de interpretabilidade:** Poucos estudos (15%) abordam explicabilidade de modelos de IA.
- **Ausência de estudos longitudinais:** A maioria dos estudos tem duração limitada (< 1 semestre).

4.7.2 Lacunas Pedagógicas

- **Desalinhamento curricular:** Necessidade de sistemas alinhados à BNCC (Base Nacional Comum Curricular) brasileira.
- **Foco limitado em metacognição:** Poucos sistemas abordam habilidades metacognitivas e autorregulação.
- **Integração com práticas docentes:** Falta de suporte para apropriação pedagógica pelos professores.

4.7.3 Direcionamento para Fase 2

Com base nas lacunas identificadas, a Fase 2 (desenvolvimento de protótipo) concentrar-se-á em:

1. **Sistema explicável:** Desenvolvimento de sistema de IA com explicabilidade integrada (XAI).
2. **Alinhamento curricular:** Adequação à BNCC, especialmente competências matemáticas do Ensino Fundamental II e Ensino Médio.
3. **Validação ecológica:** Planejamento de experimentos em contexto escolar real com grupo controle.
4. **Foco em metacognição:** Incorporação de estratégias de autorregulação e feedback metacognitivo.
5. **Dados abertos e reprodutibilidade:** Compartilhamento de código, dados e modelos treinados.

4.8 PRÓXIMOS PASSOS

A conclusão da Fase 1 (Projeto de TCC — PTC) com esta revisão sistemática estabelece fundação sólida para as fases subsequentes:

4.8.1 Fase 2 — Desenvolvimento de Protótipo (fevereiro–julho/2026)

- Levantamento de requisitos baseado nas lacunas identificadas
- Desenvolvimento de protótipo de sistema explicável
- Alinhamento com competências BNCC
- Implementação de componentes de autorregulação
- Testes unitários e de integração

4.8.2 Fase 3 — Validação Experimental (julho–novembro/2026)

- Desenho experimental com grupo controle
- Execução de experimentos em contexto escolar real
- Coleta de dados quantitativos e qualitativos
- Análise estatística de resultados
- Discussão de implicações para prática e política educacional

4.8.3 Síntese Final do TCC (outubro–novembro/2026)

- Integração de resultados das três fases
- Discussão de contribuições teóricas e práticas
- Recomendações para pesquisas futuras
- Submissão e defesa do TCC

5 CRONOGRAMA

O cronograma do projeto está organizado em três fases principais, conforme descrito no Capítulo 1, abrangendo o período de março de 2025 a novembro de 2026 (20 meses). A Tabela 6 apresenta as atividades planejadas e sua distribuição temporal.

Tabela 6 – Cronograma de atividades do projeto (março/2025 – novembro/2026).

Período	Atividade	Fase	Status
Março/2025	Definição sobre o tema com orientador/coorientador	Fase 1	Concluído
Abril/2025	Revisão bibliográfica inicial	Fase 1	Concluído
Junho/2025	Refatoração da revisão sistemática	Fase 1	Concluído
Novembro/2025	Leitura dos artigos + criação do PTC	Fase 1	Em andamento
Fevereiro/2026	Levantamento de requisitos da solução	Fase 2	Planejado
Março/2026	Desenvolvimento do protótipo	Fase 2	Planejado
Julho/2026	Execução dos experimentos	Fase 3	Planejado
Setembro/2026	Análise dos resultados	Fase 3	Planejado
Outubro/2026	Correção do texto do TCC	Fase 3	Planejado
Novembro/2026	Submissão do TCC	Fase 3	Planejado

Observações:

- **Fase 1 (PTC):** Revisão sistemática da literatura (março–novembro/2025).
- **Fase 2 (TCC):** Desenvolvimento de protótipo de ferramenta computacional para educação matemática (fevereiro–junho/2026).
- **Fase 3 (TCC):** Validação experimental do protótipo em contexto escolar real (julho–novembro/2026).

O projeto encontra-se atualmente na reta final da Fase 1, com a documentação da revisão sistemática realizada conforme protocolo PRISMA 2020 (PAGE et al., 2021). As fases subsequentes serão executadas durante o TCC propriamente dito, utilizando os resultados e lacunas identificadas nesta primeira fase como direcionamento para o desenvolvimento e validação da ferramenta educacional.

REFERÊNCIAS

APPIAH-ODAME, Eric. K. Authentic Assessment for Motivating Student Learning and Teaching Effectiveness in Rural, High-Need Secondary Schools in Manitoba, Canada. **European Journal of Mathematics and Science Education**, 2024. Relevance: 4.0/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/6a34a705ae5cbfdcfcbce84314d5edd3693cf786f>>.

CHITRE, Vidya. Exploring Machine Learning Techniques for Predictive Analytics in Computational Mathematics. **Panamerican Mathematical Journal**, 2024. Relevance: 4.1/10; Source: crossref. Disponível em: <<https://doi.org/10.52783/pmj.v34.i2.919>>.

DEPREN, Serpil Kılıç; AŞKIN Öyküm Esra; ÖZ, Ersoy. Identifying the Classification Performances of Educational Data Mining Methods: A Case Study for TIMSS. **Educational Sciences Theory & Practice**, 2017. Relevance: 4.2/10; Source: openalex. Disponível em: <<https://openalex.org/W2744614829>>.

HASIB, Khan Md et al. A Machine Learning and Explainable AI Approach for Predicting Secondary School Student Performance. In: **Computing and Communication Workshop and Conference**. [s.n.], 2022. Relevance: 4.1/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/d8baeb876681111e245f523d4771218142884a38>>.

JOSE, Bobby Chellanthara et al. Assessing the Effectiveness of Adaptive Learning Systems in K-12 Education. **International Journal of Advanced IT Research and Development**, 2024. Relevance: 4.0/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/b4cfe8f3adedc7ded70c91b5e95e6ea97bd842f9>>.

KUMAR, Mukesh et al. Analysis of Feature Selection and Data Mining Techniques to Predict Student Academic Performance. In: **2022 International Conference on Decision Aid Sciences and Applications (DASA)**. [s.n.], 2022. Relevance: 4.3/10; Source: openalex. Disponível em: <<https://openalex.org/W4225307903>>.

MACLELLAN, Christopher J. Computational Models of Human Learning: Applications for Tutor Development, Behavior Prediction, and Theory Testing. **Research Showcase @ Carnegie Mellon University (Carnegie Mellon University)**, 2017. Relevance: 4.2/10; Source: openalex. Disponível em: <<https://openalex.org/W2758976604>>.

MERTASARI, N. M. S.; SASTRI, Ni Luh Putu Pranena; PASCIMA, Ida Bagus Nyoman. Performance assessment: Improving metacognitive ability in mathematics learning. **Journal of Education and e-Learning Research**, 2023. Relevance: 4.0/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/18c8bdd1f25a4048e13f13067736719a96a57d58>>.

PAGE, Matthew J. et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, p. n71, 2021. Disponível em: <<https://doi.org/10.1136/bmj.n71>>.

PEJIC, Aleksandar; MOLCER, P. S.; GULAČI, Kristian. Math proficiency prediction in computer-based international large-scale assessments using a multi-class machine learning model. In: **Symposium on Intelligent Systems and Informatics**. [s.n.], 2021. Relevance: 4.5/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/d23946361105e17d3234d3fc61af3496d84c2a25>>.

SOKKHEY, Phauk et al. Multi-models of Educational Data Mining for Predicting Student Performance in Mathematics: A Case Study on High Schools in Cambodia. **IEIE Transactions on Smart Processing and Computing**, 2020. Relevance: 4.4/10; Source: openalex. Disponível em: <<https://openalex.org/W3038720994>>.

TJAHYADI, H. The Implementation of Educational Data Mining in Predicting Students' Academic Achievement in Mathematics at a Private Elementary School. **International Journal of Information and Education Technology**, 2025. Relevance: 4.5/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/0ce6836a1d14374ecb9bd5e3f36f970e98cb41ad>>.

USKOV, Vladimir L. et al. Machine Learning-based Predictive Analytics of Student Academic Performance in STEM Education. In: . [s.n.], 2019. Relevance: 4.2/10; Source: openalex. Disponível em: <<https://openalex.org/W2946865664>>.

ZHANG, Lei; ZHU, Weihua; FENG, Ling. Design of Personalized Learning Path Optimization Algorithm Based on Deep Learning. In: **International Conferences on Computers, Information Processing and Advanced Education**. [s.n.], 2025. Relevance: 4.2/10; Source: semantic_scholar. Disponível em: <<https://www.semanticscholar.org/paper/131a6b2c5194f51f601d09d70b7bb1ec2970b5>>.

ZHANG, Xiaohui. An Innovative Model of Higher Mathematics Curriculum Education Incorporating Artificial Intelligence Technology. **Applied Mathematics and Nonlinear Sciences**, 2024. Relevance: 4.2/10; Source: crossref. Disponível em: <<https://doi.org/10.2478/amns.2023.2.01524>>.

ÜNAL, Ferda. Data Mining for Student Performance Prediction in Education. **IntechOpen eBooks**, 2020. Relevance: 4.1/10; Source: openalex. Disponível em: <<https://openalex.org/W3015458685>>.