DISCALCULAI: A Computational Approach for the Identification of Dyscalculia in Educational Contexts

Jean Alex Firmino Pinhata[1], **Mateus Natan Roes Salgueiro**[1], **Alex Marino Gonçalves de Almeida**[1], **Rosemeiry de Castro Prado**[1]

[1] Departamento de Ciência de Dados – Faculdade de Tecnologia de Ourinhos (FATEC Ourinhos)
Avenida Vitalina Marcusso, 1400 – Campus Universitário
19910-206 – Ourinhos – SP – Brazil

{jean.pinhata,mateus.roes}@gmail.com
{alex.marino,rose.prado}@fatecourinhos.edu.br

Dyscalculia is a specific learning disorder that affects numerical processing and mathematical reasoning. It remains underrecognized in educational contexts due to limited teacher training and restricted access to specialized knowledge. This scenario contributes to the misinterpretation of students' difficulties, delaying appropriate pedagogical interventions and exacerbating educational inequalities. This study presents DISCALCULAI, a system designed to support teachers' literacy regarding dyscalculia and assist in identifying its indicators within educational environments. Its approach is grounded in Natural Language Processing (NLP) techniques, particularly Retrieval-Augmented Generation (RAG), to provide reliable and contextually relevant responses based on previously selected scientific literature. The system integrates text processing, semantic representation generation, vector indexing, and the use of Large Language Models (LLMs), enabling interaction through natural language queries. Unlike approaches based solely on generative models, the proposed architecture delivers evidence-based responses supported by retrieved information, reducing the occurrence of inaccurate outputs. To evaluate the system's effectiveness, an experiment was conducted comparing the RAG-based approach with a standalone language model using a set of 31 domain-specific queries. The results indicate statistically significant improvements in semantic similarity, groundedness, and error rate ($p = 0.0226$), demonstrating that the integration of retrieval mechanisms substantially enhances the relevance and contextual quality of the responses. By contributing to the recognition of dyscalculia indicators, DISCALCULAI promotes more inclusive educational environments and highlights the role of technology in reducing barriers to access specialized knowledge.

1. Introdução

A discalculia é um transtorno específico de aprendizagem caracterizado por dificuldades persistentes na compreensão de conceitos numéricos e na realização de operações matemáticas, impactando diretamente o desempenho acadêmico dos estudantes [García 1998, Bernardi and Stobäus 2011]. Apesar de sua relevância, esse transtorno ainda é frequentemente subdiagnosticado no contexto educacional, especialmente em cenários nos quais os professores não dispõem de formação adequada sobre suas características e manifestações [da Silva Brum and Lara 2019]. Como consequência, as dificuldades apresentadas pelos alunos são, muitas vezes, interpretadas como desinteresse ou

falta de esforço, o que compromete a adoção de estratégias pedagógicas apropriadas e dificulta a promoção de práticas educacionais inclusivas.

A literatura destaca que o conhecimento docente sobre a discalculia é um fator determinante para a identificação precoce de indícios do transtorno e para a implementação de intervenções pedagógicas eficazes [Shalev 1998, Kucian et al. 2021]. No entanto, observa-se uma lacuna significativa entre o conhecimento científico disponível e sua aplicação prática no ambiente escolar, o que evidencia a necessidade de soluções que facilitem o acesso dos professores a informações confiáveis e contextualizadas [Bastos 2016].

Paralelamente, avanços recentes em Inteligência Artificial (IA), especialmente na área de Processamento de Linguagem Natural (PLN), têm possibilitado o desenvolvimento de sistemas capazes de acessar, organizar e disponibilizar grandes volumes de informação de forma eficiente [Jurafsky and Martin 2025]. Modelos de linguagem de grande porte (Large Language Models – LLMs) apresentam elevada capacidade de geração textual, porém são suscetíveis à produção de informações imprecisas, fenômeno conhecido como alucinação. Nesse contexto, abordagens baseadas em Geração Aumentada por Recuperação (Retrieval-Augmented Generation – RAG) têm sido propostas como alternativa para aumentar a confiabilidade das respostas, ao integrar mecanismos de recuperação de informação com modelos generativos [Lewis et al. 2020].

Diante desse cenário, este trabalho propõe o desenvolvimento do sistema DISCALCULAI, uma ferramenta baseada em PLN e RAG, concebida para apoiar o letramento docente sobre a discalculia e auxiliar na identificação de seus indícios em contextos educacionais. O sistema permite que professores realizem consultas em linguagem natural sobre o transtorno, recuperando informações relevantes a partir de uma base de conhecimento composta por literatura científica e gerando respostas contextualizadas e fundamentadas.

As principais contribuições deste trabalho são a proposição de uma arquitetura baseada em RAG voltada à mediação de conhecimento científico no contexto educacional; o desenvolvimento de uma ferramenta de apoio ao reconhecimento de indícios de discalculia por professores e a realização de uma avaliação experimental que demonstra ganhos de relevância na utilização da abordagem proposta em comparação a modelos de linguagem isolados. Do ponto de vista técnico, a contribuição deste trabalho é incremental e está concentrada na integração e na aplicação de componentes existentes de PLN, recuperação semântica e geração textual ao contexto educacional da discalculia. Não se propõe um novo algoritmo de recuperação ou uma nova arquitetura de modelo de linguagem. A contribuição reside na construção, documentação e avaliação de um assistente especializado para apoiar o acesso docente ao conhecimento científico.

2. Trabalhos Relacionados

Wolf [Wolf 2023] investigou recursos tecnológicos aplicados às práticas pedagógicas para estudantes com Discalculia do Desenvolvimento. O estudo identificou jogos, aplicativos e plataformas de treinamento cognitivo, destacando limitações de infraestrutura e formação docente.

Bhushan et al. [Bhushan et al. 2024] revisaram técnicas de Inteligência Artificial aplicadas à triagem e intervenção em casos de discalculia. Os autores destacaram desa-

fios relacionados à escassez de dados públicos, proteção de dados de menores e limitada quantidade de estudos específicos.

Dhingra et al. [Dhingra et al. 2022] avaliaram a ferramenta adaptativa Mathlete com crianças diagnosticadas com discalculia. Os resultados indicaram benefícios de atividades personalizadas como complemento ao ensino tradicional.

No contexto das tecnologias de Processamento de Linguagem Natural (PLN) e dos Grandes Modelos de Linguagem (LLMs), Jurafsky e Martin [Jurafsky and Martin 2025] apresentam uma visão abrangente dos fundamentos teóricos e práticos que sustentam os sistemas modernos de compreensão e geração de linguagem natural. A obra discute a evolução das técnicas de PLN, desde abordagens estatísticas baseadas em n-gramas até arquiteturas Transformer e métodos contemporâneos de alinhamento e ajuste fino de modelos. Particularmente relevante para este trabalho é a discussão sobre mecanismos de recuperação de informações e geração fundamentada em evidências, que servem de base conceitual para arquiteturas Retrieval-Augmented Generation (RAG). Nesse contexto, os autores demonstram como a integração entre recuperação semântica e geração textual pode reduzir inconsistências, ampliar a confiabilidade das respostas e melhorar a utilização de conhecimento especializado. Essas contribuições fundamentam diretamente a proposta do DISCALCULAI, especialmente no uso de embeddings semânticos, indexação vetorial e geração de respostas contextualizadas para apoiar professores no acesso a informações científicas sobre discalculia.

Campos [Campos 2025] desenvolveu o ParkinsonBot, protótipo de chatbot educacional destinado ao apoio de pessoas responsáveis pela capacitação de cuidadores de idosos com doença de Parkinson. O trabalho é relevante para a presente proposta por discutir acessibilidade e adequação da linguagem ao público-alvo. O solicita respostas claras e objetivas em seu prompt, mas ainda não possui um módulo específico de simplificação textual capaz de controlar a complexidade e preservar o significado.

Os trabalhos relacionados concentram-se principalmente em diagnóstico, treinamento cognitivo e adaptação de conteúdos.Nesse sentido, o presente trabalho busca contribuir para o avanço da área por meio da integração de técnicas de PLN, recuperação semântica e modelos generativos, oferecendo uma ferramenta capaz de disponibilizar conhecimento científico sobre discalculia de forma acessível e contextualizada aos profissionais da educação.

3. Metodologia

O presente trabalho caracteriza-se como uma pesquisa aplicada orientada ao desenvolvimento de um sistema inteligente de apoio à tomada de decisão pedagógica no contexto da discalculia. A proposta fundamenta-se na integração entre técnicas de Processamento de Linguagem Natural (PLN), recuperação semântica de informações e Grandes Modelos de Linguagem (Large Language Models – LLMs), empregando a arquitetura Retrieval-Augmented Generation (RAG). Segundo Lewis et al. [Lewis et al. 2020], a arquitetura RAG combina mecanismos de recuperação de conhecimento externo com modelos generativos, permitindo reduzir limitações associadas ao conhecimento paramétrico dos LLMs e aumentar a confiabilidade das respostas produzidas.

A ferramenta permite que professores e outros profissionais da educação realizem consultas em linguagem natural sobre o transtorno, como, por exemplo, sintomas, carac-

terísticas e dificuldades associadas. A partir dessas consultas, o sistema retorna respostas estruturadas, objetivas e fundamentadas em artigos científicos previamente selecionados.

A metodologia adotada foi organizada em cinco etapas principais: construção da base de conhecimento, pré-processamento textual, indexação vetorial, recuperação semântica de informações e geração de respostas fundamentadas em evidências documentais.

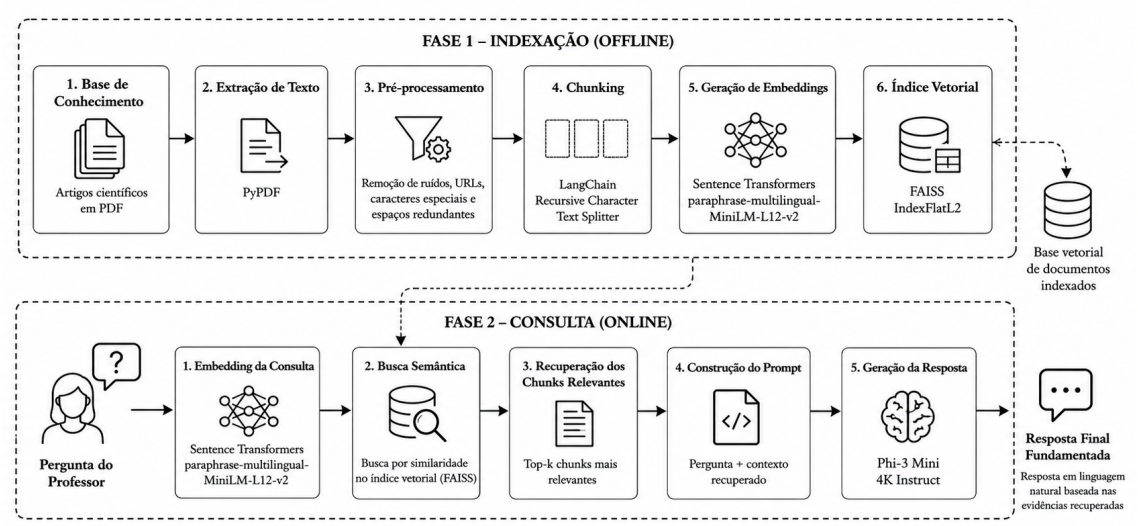


Figura 1. Arquitetura geral do sistema DISCALCULAI baseada em Retrieval-Augmented Generation (RAG).

A primeira etapa consistiu na construção da base documental do sistema. A base utilizada pelo sistema contém 30 documentos em formato PDF, sendo artigos científicos e materiais especializados relacionados à discalculia, dificuldades de aprendizagem matemática e estratégias educacionais inclusivas. Os documentos foram armazenados em formato PDF e submetidos a um processo automatizado de extração textual utilizando a biblioteca PyPDF. A utilização de fontes científicas especializadas visa garantir a qualidade e a confiabilidade das informações disponibilizadas pelo sistema, aspecto considerado fundamental em aplicações educacionais e de suporte à tomada de decisão.

A busca documental foi realizada no Google Acadêmico, no período de 2011 a 2025, utilizando os termos ”Discalculia”e ”Discalculia do Desenvolvimento”. Foram incluídos trabalhos com texto integral disponível e relação direta com os temas definidos. Foram excluídos documentos fora do domínio, sem texto extraível ou com conteúdo insuficiente. As versões duplicadas foram removidas antes da indexação.

Após a extração do conteúdo textual, foi realizada uma etapa de pré-processamento dos documentos. Essa etapa incluiu a remoção de caracteres especiais, espaços redundantes, URLs e elementos que não contribuíam para a representação semântica do conteúdo. De acordo com Jurafsky e Martin [Jurafsky and Martin 2025], procedimentos de normalização textual contribuem para a redução de ruídos linguísticos e favorecem a construção de representações semânticas mais consistentes, melhorando o desempenho das etapas subsequentes de recuperação e geração.

Na sequência, os documentos foram segmentados em unidades textuais menores,

denominadas *chunks*, por meio do algoritmo Recursive Character Text Splitter, disponibilizado pela biblioteca LangChain. A segmentação documental constitui uma prática amplamente empregada em sistemas RAG, uma vez que modelos de recuperação operam de maneira mais eficiente sobre fragmentos menores de texto do que sobre documentos completos [Gao et al. 2023]. Além disso, a utilização de sobreposição entre *chunks* permite preservar relações contextuais entre segmentos adjacentes, reduzindo perdas semânticas decorrentes da fragmentação documental.

O Recursive Character Text Splitter foi configurado com tamanho máximo de 1.200 caracteres e sobreposição de 250 caracteres. Para cada consulta, foram recuperados os três fragmentos de menor distância no índice FAISS IndexFlatL2. Não foi aplicado limiar mínimo de relevância nem procedimento adicional de reranqueamento. O modelo de embeddings utilizado foi o paraphrase-multilingual-MiniLM-L12-v2, que produz vetores de 384 dimensões.

Para representar semanticamente cada fragmento textual, foi utilizado o modelo Sentence Transformers *paraphrase-multilingual-MiniLM-L12-v2*. Segundo Reimers e Gurevych [Reimers and Gurevych 2019], modelos Sentence-BERT são capazes de gerar embeddings semânticos densos que preservam relações de significado entre sentenças e documentos, permitindo a comparação eficiente de similaridade textual por meio de operações vetoriais. A escolha de um modelo multilíngue justifica-se pela necessidade de representar adequadamente conteúdos em língua portuguesa, mantendo simultaneamente eficiência computacional e qualidade semântica.

Os embeddings gerados foram armazenados em uma estrutura de indexação vetorial utilizando a biblioteca FAISS (Facebook AI Similarity Search). Conforme demonstrado por Johnson, Douze e Jégou [Johnson et al. 2019], o FAISS constitui uma das principais soluções para busca aproximada por similaridade em espaços vetoriais de alta dimensionalidade, possibilitando recuperação eficiente mesmo em coleções documentais extensas. Nesta pesquisa, foi empregada a estrutura *IndexFlatL2*, que realiza buscas baseadas na distância euclidiana entre vetores.

Durante a execução do sistema, as consultas formuladas pelos usuários são inicialmente convertidas para o mesmo espaço vetorial utilizado na indexação dos documentos. Em seguida, realiza-se uma busca semântica no índice FAISS, recuperando os fragmentos documentais mais relevantes para a consulta. Esse mecanismo permite que o sistema acesse conhecimento especializado armazenado externamente ao modelo de linguagem, reduzindo a dependência exclusiva dos parâmetros aprendidos durante o treinamento do LLM.

A etapa de geração textual foi implementada utilizando o modelo Phi-3 Mini 4K Instruct, desenvolvido pela Microsoft. O modelo foi selecionado em função de sua relação entre desempenho e eficiência computacional, característica importante para aplicações acadêmicas executadas em ambientes com recursos limitados. Conforme discutido por Jurafsky e Martin [Jurafsky and Martin 2025], modelos instruídos (*instruction-tuned models*) apresentam maior capacidade de seguir orientações explícitas presentes em prompts estruturados, favorecendo a geração de respostas mais alinhadas às necessidades da tarefa.

Não foi realizado ajuste fino do Phi-3 ou do modelo de embeddings. Os documentos foram utilizados exclusivamente como base externa de recuperação durante a in-

ferência. As 31 perguntas foram empregadas somente na avaliação experimental e não constituem um conjunto de treinamento.

A arquitetura RAG operou por meio da incorporação dos fragmentos recuperados ao prompt enviado ao modelo de linguagem. Dessa forma, o LLM recebe simultaneamente a consulta do usuário e o contexto documental considerado relevante, sendo instruído a gerar respostas fundamentadas exclusivamente nas evidências fornecidas. Essa estratégia segue as recomendações apresentadas por Lewis et al. [Lewis et al. 2020] e Gao et al. [Gao et al. 2023], que destacam a recuperação documental como mecanismo eficaz para redução de alucinações e aumento da precisão factual em sistemas baseados em LLMs.

Por último, foi implementada uma etapa de pós-processamento destinada à remoção de fragmentos incompletos, repetições excessivas e artefatos decorrentes da geração textual. O resultado final consiste em respostas produzidas em linguagem natural, fundamentadas em literatura científica especializada e direcionadas ao apoio de professores que atuam junto a estudantes com discalculia.

Destaca-se que a ferramenta não realiza diagnóstico, mas atua como um sistema de apoio à tomada de decisão pedagógica, auxiliando na observação e interpretação de possíveis sinais de discalculia.

Espera-se, como resultado da utilização do sistema, promover avanços no processo de aprendizagem matemática, bem como contribuir para a redução de desigualdades educacionais associadas à ausência de diagnóstico e de suporte pedagógico adequado para estudantes com transtornos de aprendizagem.

4. Avaliação Experimental

O objetivo desta avaliação é analisar a qualidade das respostas geradas pelo sistema DISCALCULAI e verificar os ganhos proporcionados pela utilização da abordagem baseada em Geração Aumentada por Recuperação (RAG).

4.1. Configuração do Experimento

O conjunto de avaliação foi composto por 31 perguntas relacionadas ao domínio da discalculia, contemplando aspectos conceituais, diagnósticos, cognitivos e pedagógicos frequentemente abordados na literatura especializada. As mesmas perguntas foram submetidas aos dois cenários experimentais, possibilitando comparação direta entre as respostas produzidas.

O conjunto de avaliação foi definido antes da execução dos experimentos e mantido idêntico nos dois cenários. As 31 perguntas foram distribuídas em quatro categorias: conceitos, cognição e etiologia (6 perguntas); indicadores, manifestações e impactos (10); ação docente, intervenção e encaminhamento (10); e IA, PLN, RAG e letramento tecnológico (5).

Todas as respostas analisadas foram geradas automaticamente pelo Phi-3 Mini. Cada uma das 31 perguntas foi submetida ao modelo em dois cenários: sem acesso à base documental e com recuperação RAG. Nenhuma dessas respostas foi redigida ou corrigida pelos autores antes do cálculo das métricas.

A avaliação foi estruturada em duas dimensões complementares: desempenho operacional e qualidade semântica das respostas. Essa organização segue as recomendações de [Gao et al. 2023], que destacam a necessidade de avaliar sistemas RAG não apenas quanto à eficiência computacional, mas também em relação à qualidade das informações recuperadas e utilizadas durante a geração textual.

4.2. Avaliação Operacional

A primeira dimensão avaliada refere-se ao desempenho operacional do sistema. Foram consideradas as seguintes métricas:

- **Tempo Médio de Resposta:** mede o intervalo necessário para gerar uma resposta completa a partir de uma consulta do usuário, permitindo quantificar o impacto computacional introduzido pela etapa de recuperação semântica.
- **Quantidade Média de Tokens:** representa o tamanho médio das respostas produzidas, auxiliando na identificação de possíveis alterações no comportamento de geração do modelo quando submetido ao contexto recuperado.

4.3. Avaliação da Qualidade Semântica

A segunda dimensão contempla métricas relacionadas à qualidade das respostas geradas.

A **Similaridade Semântica Média** foi calculada por meio da similaridade do cosseno entre embeddings das respostas geradas e respostas de referência previamente definidas para um subconjunto de perguntas. Os embeddings foram obtidos através do modelo *paraphrase-multilingual-MiniLM-L12-v2*. Conforme demonstrado por [Reimers and Gurevych 2019], essa abordagem permite mensurar a proximidade semântica entre textos mesmo na ausência de correspondência lexical exata. A similaridade entre embeddings mensura proximidade semântica, mas não comprova correção fatual, duas respostas podem apresentar elevada similaridade vetorial e, ainda assim, expressar informações diferentes ou contraditórias.

O **Groundedness Médio** foi utilizado para avaliar o grau de aderência das respostas ao contexto documental recuperado. Para cada consulta, calculou-se a similaridade de cosseno entre os embeddings da resposta gerada e os embeddings do contexto utilizado durante a geração. Segundo [Gao et al. 2023], métricas de groundedness são particularmente relevantes em sistemas RAG por indicarem em que medida o modelo efetivamente ancora suas respostas nas evidências recuperadas.

A **Taxa Média de Repetição** foi empregada para identificar redundâncias textuais nas respostas geradas, sendo calculada como o complemento da proporção de palavras únicas em relação ao total de palavras da resposta. Valores mais elevados indicam maior repetição lexical e menor diversidade textual.

A **Taxa Média de Erro** funcionou como indicador da ocorrência de respostas inadequadas ou inconsistentes, identificadas por meio da presença de padrões indesejados — como termos em inglês, marcações estruturais do modelo ou fragmentos fora do domínio esperado. Embora não constitua uma métrica padronizada na literatura de avaliação de sistemas RAG, sua inclusão permitiu capturar comportamentos problemáticos observados empiricamente durante a execução do sistema.

4.4. Análise Estatística

Além da análise descritiva, foi aplicado o teste t de Student para comparação entre os resultados obtidos nos dois cenários experimentais. O teste foi aplicado sobre os valores de similaridade semântica obtidos para as 31 perguntas avaliadas, resultando em $p = 0{,}0226$. Esse resultado indica que as diferenças observadas entre o modelo puro e o modelo com RAG são estatisticamente significativas ao nível de 5%, reforçando a hipótese de que a incorporação do mecanismo de recuperação semântica produz ganhos relevantes na qualidade informacional das respostas geradas.

A comparação entre o modelo com e sem RAG foi utilizada para observar o efeito da inclusão do contexto recuperado. Entretanto, o cenário com RAG possui acesso a informações externas indisponíveis ao modelo isolado. Assim, os resultados não demonstram superioridade sobre todas as estratégias de especialização. Trabalhos futuros deverão incluir modelos adaptados por LoRA, QLoRA ou técnicas equivalentes, utilizando conjuntos independentes de treinamento, validação e teste [Hu et al. 2022, Dettmers et al. 2023].

5. Resultados

Os resultados foram organizados em tabela e gráficos comparativos, permitindo analisar os impactos da recuperação semântica tanto sob a perspectiva operacional quanto informacional. Os exemplos apresentados correspondem às saídas originais geradas pelo modelo durante o experimento, sem edição manual de seu conteúdo.

5.1. Resultados Quantitativos

A Tabela 1 apresenta os resultados obtidos na comparação entre o modelo de linguagem utilizado isoladamente (LLM Puro) e sua versão integrada à arquitetura RAG (LLM + RAG).

Tabela 1. Comparação quantitativa entre o modelo puro e o modelo com RAG.

Métrica	LLM Puro	LLM + RAG
Tempo Médio de Resposta (s)	8,1840	9,2613
Quantidade Média de Tokens	70,7097	72,8065
Taxa Média de Repetição	0,0802	0,2468
Similaridade Semântica Média	0,5332	0,6420
Taxa Média de Erro	0,0645	0,0000
Groundedness Médio	0,6086	0,6817

A incorporação da arquitetura RAG produziu um aumento discreto no tempo médio de resposta, de 8,18 s para 9,26 s, acréscimo esperado, dado que a recuperação semântica adiciona etapas de geração de embeddings da consulta e busca vetorial no índice FAISS. A diferença aproximada de 1,08 segundos indica que o mecanismo de recuperação não compromete a usabilidade do sistema. A Figura 2 apresenta as métricas de desempenho operacional.

Quanto ao tamanho das respostas, não foram observadas diferenças substanciais entre os cenários: o modelo com RAG produziu respostas ligeiramente mais concisas, mantendo nível de detalhamento comparável ao do modelo puro.

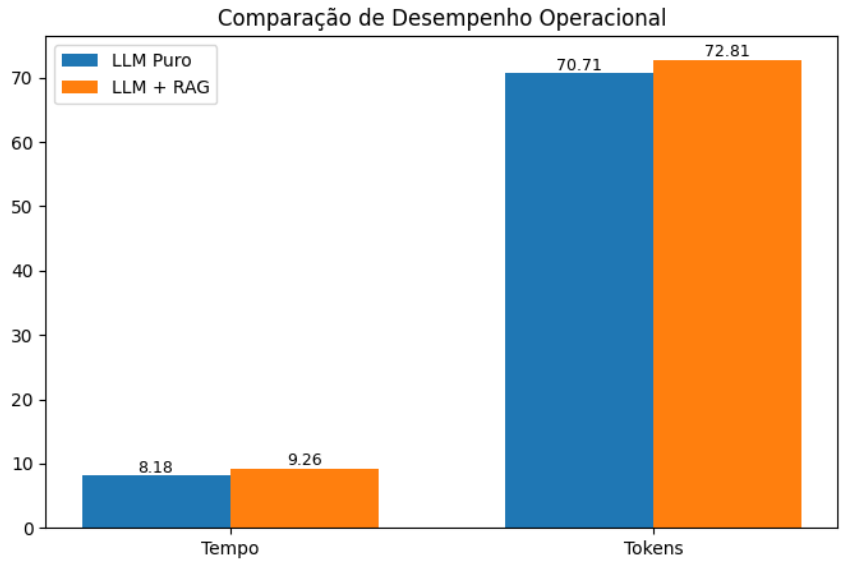


Figura 2. Comparação do desempenho operacional entre o modelo puro e o modelo com RAG.

As melhorias mais expressivas concentraram-se nas métricas de qualidade informacional. A Similaridade Semântica Média aumentou de 0,5332 para 0,6420, representando ganho aproximado de 20,4%, o que indica maior alinhamento das respostas ao conteúdo esperado sobre discalculia. O Groundedness Médio cresceu de 0,6086 para 0,6817, evidenciando que o modelo passou a ancorar suas respostas de forma mais efetiva nas evidências recuperadas. Essa diferença mostrou-se estatisticamente significativa segundo o teste t de Student ($p = 0{,}0226$), reforçando a evidência de que a incorporação do mecanismo RAG contribuiu para o aumento da qualidade semântica das respostas.

A Taxa Média de Erro caiu de 0,0645 para 0,0000 no cenário com RAG, sugerindo que o acesso à base documental contribuiu para eliminar respostas inadequadas ou fora do domínio esperado.

Por outro lado, a Taxa Média de Repetição elevou-se de 0,0802 para 0,2468. Esse comportamento pode ser atribuído à incorporação de fragmentos documentais com terminologia recorrente ao prompt do modelo, induzindo maior reuso lexical nas respostas geradas. Ainda assim, os ganhos em Similaridade Semântica, Groundedness e Taxa de Erro indicam que esse efeito não comprometeu a qualidade geral das respostas.

As Figuras 2 e 3 complementam a análise quantitativa, evidenciando que os ganhos proporcionados pela arquitetura RAG concentram-se principalmente na dimensão informacional, enquanto os custos computacionais adicionais permanecem baixos.

5.2. Análise Qualitativa das Respostas

Além da avaliação quantitativa, foi realizada uma análise qualitativa com o objetivo de ilustrar as diferenças entre os dois cenários avaliados. A seguir, apresenta-se um exemplo de resposta produzida pela arquitetura RAG e um exemplo produzido pelo modelo sem acesso à base documental.

Resposta produzida pela arquitetura RAG

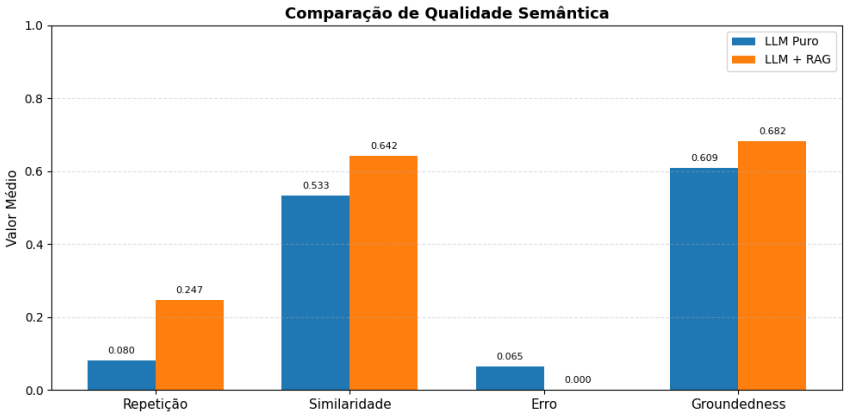


Figura 3. Comparação da qualidade semântica entre o modelo puro e o modelo com RAG.

Pergunta: Quais áreas do cérebro estão associadas à discalculia?

Resposta:

1. O lobo frontal, responsável pela concentração, planejamento, iniciativa e cálculos mentais rápidos.
2. O lobo parietal esquerdo, que processa informações relacionadas às noções de espaço e volume.
3. O lobo occipital, que é o centro da visão e diferencia objetos de cores e texturas semelhantes.
4. O lobo temporal, responsável pela percepção auditiva, memória verbal em longo prazo e realizações matemáticas básicas.

Nesse exemplo, a resposta apresenta informações específicas sobre a neurobiologia da aprendizagem matemática: além de identificar as regiões cerebrais envolvidas, o sistema descreve as funções cognitivas associadas a cada estrutura, demonstrando utilização efetiva do conhecimento recuperado da base documental.

Resposta produzida pelo modelo sem RAG

Pergunta: Quais são os sintomas da discalculia?

Resposta:

1. Dificuldade para entender conceitos básicos de aritmética, como adição, subtração, multiplicação e divisão.
2. Problemas com a memorização das regras matemáticas fundamentais.
3. Tempo prolongado na resolução de problemas simples que requerem raciocínio lógico;
4. Confusão entre números semelhantes (por exemplo, confundir o número dez com vinte);
5. Falta de motivação por tarefas relacionadas às disciplinas matemáticas;
6. Desenvolvimento atrasado no aprendizado de habilidades numéricas essenciais para outras áreas do currículo escolar.

Embora a resposta seja coerente e contenha informações corretas, seu caráter é notadamente mais genérico. Os conceitos são descritos de forma ampla, sem referências a evidências específicas ou aprofundamento conceitual característico da literatura especializada. De maneira geral, a análise qualitativa reforça que a arquitetura RAG não apenas melhora os indicadores quantitativos, mas também produz respostas mais ricas em conteúdo especializado, diferença que, no contexto educacional, pode representar um suporte mais efetivo ao trabalho docente com estudantes com discalculia.

Não foi realizada validação humana por professores ou especialistas. Portanto, os resultados não permitem afirmar que as respostas sejam fatualmente corretas, pedagogicamente úteis ou seguras. As métricas apresentadas descrevem somente o comportamento automático do protótipo e não demonstram impacto educacional.

6. Conclusão

Este trabalho apresentou o DISCALCULAI, um sistema baseado em técnicas de Processamento de Linguagem Natural e Geração Aumentada por Recuperação (RAG), desenvolvido com o objetivo de apoiar o letramento docente sobre a discalculia e auxiliar na identificação de seus indícios em contextos educacionais. A proposta buscou enfrentar a dificuldade de acesso a conhecimento especializado por parte dos professores, fator que frequentemente contribui para a interpretação inadequada das dificuldades de aprendizagem em matemática. A arquitetura desenvolvida integra etapas de processamento textual, geração de representações semânticas e recuperação de informações em bases científicas, permitindo a construção de respostas fundamentadas e contextualizadas. A complexidade textual das respostas constitui uma limitação adicional. Termos técnicos e explicações extensas podem dificultar o uso por professores não especialistas. Como trabalho futuro, pretende-se incorporar um módulo de simplificação capaz de produzir versões técnica e acessível da resposta, preservando os conceitos essenciais, porém buscando gerar respostas mais simplificadas aos usuários. Os resultados experimentais indicaram que a abordagem baseada em RAG apresenta desempenho superior em termos de relevância e qualidade das respostas quando comparada a modelos de linguagem isolados, evidenciando a importância da incorporação de mecanismos de recuperação de informação em sistemas baseados em linguagem natural. Do ponto de vista educacional, o sistema demonstra potencial como ferramenta de apoio à prática docente, contribuindo para o reconhecimento de possíveis indícios de discalculia e favorecendo o encaminhamento adequado dos estudantes a profissionais especializados. Dessa forma, a proposta não apenas amplia o acesso ao conhecimento científico, mas também contribui para a promoção de práticas educacionais mais inclusivas. Como trabalhos futuros, sugere-se a ampliação da base de conhecimento com novos estudos científicos, a realização de avaliações com usuários em contexto real de sala de aula e a investigação de mecanismos adicionais de personalização das respostas. Além disso, pretende-se explorar a aplicação da abordagem proposta em outros transtornos de aprendizagem, ampliando o impacto da solução no contexto educacional.

Referências

Bastos, J. A. (2016). *Discalculia: Transtorno Específico da Aprendizagem em Matemática*. InterSaberes, Curitiba.

Bernardi, J. and Stobäus, C. (2011). Discalculia: conhecer para incluir. *Revista Educação Especial*, 24(39):47–60.

Bhushan, S., Arunkumar, S., Eisa, T. A. E., Nasser, M., Singh, A. K., and Kumar, P. (2024). Ai-enhanced dyscalculia screening: A survey of methods and applications for children. *Diagnostics*, 14(13):1441.

Campos, M. N. (2025). Parkinsonbot: chatbot para auxílio e acessibilidade sobre a doença de parkinson. Trabalho de conclusão de curso, Graduação em Ciência da Computação. Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre.

da Silva Brum, E. and Lara, I. C. M. (2019). Discalculia: um estudo sobre as dificuldades de aprendizagem em matemática. *Revista de Educação*, 44(2):1–15.

Dettmers, T., Pagnoni, A., Holtzman, A., and Zettlemoyer, L. (2023). Qlora: Efficient finetuning of quantized llms. In *Advances in Neural Information Processing Systems*, volume 36, pages 10088–10115.

Dhingra, K., Aggarwal, R., Garg, A., Pujari, J., and Yadav, D. (2022). Mathlete: An adaptive assistive technology tool for children with dyscalculia. *Disability and Rehabilitation: Assistive Technology*, 19(1):9–15.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., and Wang, M. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.

García, J. N. (1998). *Manual de Dificuldades de Aprendizagem: Linguagem, Leitura, Escrita e Matemática*. Artes Médicas, Porto Alegre.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., and Chen, W. (2022). Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

Johnson, J., Douze, M., and Jégou, H. (2019). Billion-scale similarity search with gpus. *IEEE Transactions on Big Data*, 7(3):535–547.

Jurafsky, D. and Martin, J. H. (2025). Speech and language processing. `https://web.stanford.edu/~jurafsky/slp3/`. 3rd Edition Draft.

Kucian, K., McCaskey, U., von Aster, M., and Tuura, R. O. (2021). Dyscalculia: A neurocognitive perspective. *Developmental Medicine & Child Neurology*, 63(4):379–385.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., and Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. In *Advances in Neural Information Processing Systems*, volume 33, pages 9459–9474.

Reimers, N. and Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of EMNLP-IJCNLP*, pages 3982–3992.

Shalev, R. (1998). Developmental dyscalculia. *Journal of Child Neurology*, 13(6):281–286.

Wolf, A. (2023). Tecnologias nas metodologias docentes para alunos com discalculia do desenvolvimento. *Educação Matemática Sem Fronteiras*, 5:76–88.