

Uso de IA no ensino e na formação profissional: Proposta de ferramenta para avaliação interativa de casos clínicos

Thiago Ferreira Leal, Érico Amaral, Julio Saraçol Domingues Júnior

¹Universidade Federal do Pampa (UNIPAMPA)

{thiagoleal.aluno, ericoamaral, juliiodomingues}@unipampa.edu.br

Abstract. *The use of artificial intelligence in health education has expanded the possibilities for simulation and assessment, but it still often occurs in a manual, poorly standardized, and low-traceability manner. This paper presents EducaSim, a web platform for applying interactive clinical-case assessments with support from LLMs, based on textual references, web pages, and PDF files. The study is applied and exploratory. Preliminary results indicate feasibility in multimodal ingestion, reference-grounded case generation, learning-theory guidance, and the production of interactive reports.*

Resumo. *O uso de inteligência artificial na educação em saúde tem ampliado as possibilidades de simulação e avaliação, mas ainda ocorre, em muitos casos, de forma manual, pouco padronizada e com baixa rastreabilidade. Este artigo apresenta o EducaSim, uma plataforma web voltada à aplicação de avaliações interativas de casos clínicos com apoio de LLMs, a partir de referências textuais, páginas web e arquivos PDF. O estudo é aplicado e exploratório. Os resultados preliminares indicam viabilidade na ingestão multimodal, na geração de casos ancorados nas referências, na orientação por teoria de aprendizagem e na produção de relatórios interativos.*

1. Introdução

O uso de LLMs (*Large Language Models*) na área de saúde cresceu rapidamente com a popularização de interfaces conversacionais como o ChatGPT [Aster et al. 2024]. No campo educacional, essas ferramentas passaram a ser usadas para construir explicações, gerar perguntas e conduzir atividades formativas e de simulação em diferentes contextos de capacitação [Abadam-Eremeev et al. 2025, Guo et al. 2025, Loubbairi et al. 2025]. Entretanto, quando a interação ocorre de forma manual, com *prompts* copiados entre alunos e sem mediação sistêmica, surgem limitações relevantes, como baixo controle sobre o conteúdo aplicado, dificuldade para padronizar a experiência entre turmas, pouca rastreabilidade do processo e ausência de evidências estruturadas para supervisão docente.

Esse problema motivou o desenvolvimento do projeto EducaSim. O ponto de partida foi uma avaliação interativa baseada em casos clínicos na área de fisioterapia respiratória, inicialmente conduzida com ChatGPT a partir de ementas e materiais de referência definidos pelo docente. Embora a estratégia tenha mostrado potencial para criar diálogos ricos e *feedback* imediato, ela dependia da ação manual de cada estudante e não oferecia um fluxo confiável para transformar referências em avaliações reaplicáveis, auditáveis e alinhadas ao objetivo de aprendizagem.

Diante desse cenário, este trabalho apresenta o EducaSim como uma plataforma *web* para avaliação interativa de casos clínicos com apoio de IA (Inteligência Artificial). A proposta central é substituir o uso *ad hoc* de LLMs por um *pipeline* que recebe referências em texto, URLs (*Uniform Resource Locators*) e arquivos PDF (*Portable Document Format*), consolida esse material em uma base de contexto estruturada, extrai temas e competências, planeja os casos segundo uma teoria de aprendizagem selecionada, gera os casos clínicos com um modelo principal, valida o resultado com um modelo secundário e, por fim, conduz a sessão interativa com rastreabilidade e relatório docente. O sistema implementa opções de personalização da experiência do estudante, permitindo selecionar a teoria de aprendizagem que orientará a simulação; na versão atual, essa camada é inicialmente operacionalizada com base na teoria de aprendizagem experiencial de Kolb.

2. Materiais e métodos

Este trabalho caracteriza-se como pesquisa aplicada, de natureza exploratória e descritiva, orientada ao desenvolvimento de uma solução computacional para apoio à formação em saúde. Os materiais analisados incluíram uma atividade real de avaliação interativa mediada por ChatGPT na área de fisioterapia cardiopulmonar e pneumologia, artefatos do MVP (*minimum viable product*) do EducaSim, registros de funcionamento do fluxo de avaliação e estudos recuperados na literatura sobre LLMs em educação em saúde. A atividade inicial permitiu identificar requisitos como controle das fontes, progressão pedagógica, registro do desempenho e produção de evidências para acompanhamento docente.

A revisão da literatura buscou responder a duas questões: como LLMs vêm sendo usados em intervenções educacionais estruturadas em saúde e em que medida essas intervenções explicitam teorias ou estruturas pedagógicas. O protocolo adotou recorte temporal de 2022 a 2025 e combinou buscas em PubMed, IEEE Xplore, ACM Digital Library, Scopus, Google Scholar como fonte suplementar e consultas a fontes brasileiras indexadas na SBC OpenLib. As buscas foram adaptadas por base, mantendo termos relacionados a LLMs, educação, saúde, simulação, avaliação e desenho pedagógico.

Foram identificados 36 registros, reduzidos a 35 após deduplicação. Na triagem por título e resumo, 19 estudos seguiram para avaliação em texto completo, dos quais 14 foram incluídos na síntese: 10 diretamente relacionados à educação em saúde e 4 tangenciais, usados para contextualizar tendências em IA generativa e educação. A extração considerou ferramenta ou modelo utilizado, contexto educacional, tipo de estudo, presença de estrutura pedagógica, achados relevantes e limitações. Como limitação metodológica, a triagem e a extração foram conduzidas por um único revisor.

A inspeção técnico-funcional avaliou o fluxo de criação da atividade, ingestão de referências textuais, *web* e em PDF, consolidação do contexto, extração temática, planejamento pedagógico, geração e validação de casos, condução da avaliação interativa, persistência das interações e geração de relatório. Como o projeto ainda não passou por estudo com participantes, a validação apresentada neste artigo permanece restrita à viabilidade funcional do artefato, sem medir aprendizagem, retenção, transferência de conhecimento ou comparação experimental com outras abordagens.

3. Trabalhos correlatos

O uso de LLMs em educação em saúde já aparece na literatura em pelo menos três frentes recorrentes: apoio instrucional, avaliação e simulação. Estudos recentes mostram aplicações para correção de respostas abertas, tutoria e organização de atividades formativas, enquanto revisões de literatura destacam o crescimento acelerado do tema e a necessidade de desenhos mais robustos para uso educacional [Abadam-Eremeev et al. 2025, Aster et al. 2024]. No eixo de simulação, plataformas como OSCEai mostram que LLMs podem sustentar interações clínicas estruturadas, com foco em comunicação, raciocínio clínico e tomada de decisão por meio de interações guiadas [Guo et al. 2025]. Revisões mais amplas confirmam que pacientes virtuais baseados em LLMs se tornaram um dos usos mais promissores do campo, embora ainda com desafios de realismo, consistência e avaliação padronizada [Loubbairi et al. 2025].

Além da camada conversacional, cresce o interesse por arquiteturas capazes de ancorar respostas em bases de conhecimento, recuperar evidências e submeter saídas a rotinas de crítica e refinamento. Em contexto médico, fluxos do tipo RAG (*retrieval-augmented generation*) têm mostrado ganhos de precisão, utilidade e segurança quando o modelo passa a operar sobre coleções especializadas em vez de depender apenas de conhecimento generalista [Luo et al. 2024]. Para simulação clínica, arquiteturas baseadas em recuperação, raciocínio e geração também vêm sendo exploradas para reduzir alucinações e manter fidelidade dos casos clínicos [Yu et al. 2025]. Na literatura de IA generativa, propostas como Self-RAG e Self-Refine reforçam a utilidade de ciclos explícitos de recuperação, crítica e regeneração para elevar a veracidade e a qualidade das saídas [Asai et al. 2023, Madaan et al. 2023].

No contexto brasileiro, também é relevante notar a presença de iniciativas voltadas à construção de casos clínicos e pacientes virtuais, como a ferramenta apresentada por Flores et al. [Flores et al. 2023]. Embora a literatura recente demonstre avanços no uso de LLMs para simulação clínica e apoio educacional, ainda é menos frequente a proposição de plataformas que integrem, em um mesmo fluxo, a condução da avaliação interativa, o registro do processo e mecanismos de acompanhamento docente. O EducaSim dialoga com esse contexto ao propor uma solução voltada não apenas à criação do caso, mas também à organização da interação e ao acompanhamento posterior.

4. EducaSim: Plataforma proposta

O EducaSim foi concebido para transformar avaliações clínicas mediadas por IA em um processo mais controlado, reprodutível e acompanhável. A plataforma oferece um fluxo próprio para configuração, geração, validação, execução e análise da atividade com o objetivo de substituir o uso de *prompts* copiados manualmente em um *chatbot* genérico. O usuário informa o objetivo de aprendizagem, seleciona a teoria de aprendizagem aplicada e fornece materiais de referência em texto, URLs e arquivos PDF, e o sistema então opera em etapas encadeadas.

Entrada de dados. O EducaSim recebe objetivo, teoria de aprendizagem e referências em texto livre, URLs e documentos. **Ingestão e normalização.** O sistema separa os tipos de entrada, remove duplicidades, busca conteúdo *web*, segmenta documentos e consolida tudo em uma base de contexto indexada. Essa base sustenta uma estratégia de RAG (*retrieval-augmented generation*), na qual trechos relevantes são recuperados

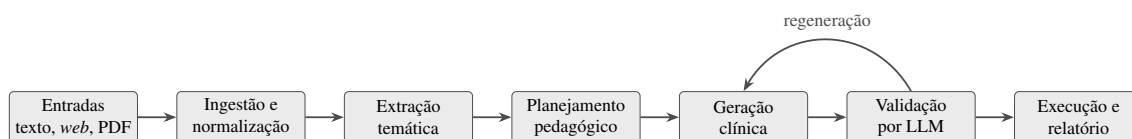


Figura 1. Pipeline do EducaSim: das referências à geração validada de casos clínicos e ao relatório docente.

para apoiar a geração, a auditoria e a tutoria. **Extração semântica.** Sobre essa base, a plataforma identifica temas, subtópicos, competências, conceitos-chave, protocolos e nível de dificuldade. **Planejamento pedagógico.** Com base nesses elementos, o sistema produz um plano instrucional dos casos, definindo progressão, número de interações, pistas disponíveis, critérios de avaliação e evidências esperadas.

Após a consolidação das referências, o EducaSim organiza a atividade em etapas de geração, auditoria e tutoria, com chamadas a LLMs realizadas via OpenRouter. Essa escolha permite empregar modelos em papéis distintos no mesmo *pipeline*, por exemplo separando a geração dos casos da revisão crítica das saídas, além de viabilizar ajustes de modelo conforme custo, janela de contexto e capacidade de seguir instruções estruturadas. A geração clínica elabora os casos, rubricas e pontos de *feedback*; a auditoria verifica aderência às fontes, consistência clínica, cobertura temática, alinhamento ao objetivo, coerência com a teoria selecionada e completude das rubricas. Quando identifica falhas, o item retorna para regeneração com indicação do motivo; apenas casos aprovados alimentam a sessão interativa e o relatório docente.

Nesse desenho, a teoria de aprendizagem opera como política pedagógica da atividade, definindo progressão, intervenção do agente, formato das perguntas, níveis de ajuda e critérios de *feedback*. Na configuração inicial, o ciclo experiencial de Kolb [Kolb 1984] é operacionalizado por quatro movimentos: apresentação de uma situação clínica concreta; solicitação de observação reflexiva, na qual o estudante explicita hipóteses e justificativas; consolidação conceitual por *feedback* ancorado nas referências; e nova aplicação do raciocínio por complementação, revisão ou caso seguinte. Esses movimentos orientam parâmetros do plano de avaliação, como ordem dos casos, critérios de avanço, política de pistas, evidências esperadas e condições para encerrar ou repetir uma etapa. A arquitetura admite a futura incorporação de outras teorias e estratégias didáticas, mas a versão atual utiliza Kolb como organização pedagógica principal.

Outro aspecto relevante é a rastreabilidade. A plataforma registra progresso por caso, consolida pontuação e gera um relatório final estruturado com resumo do desempenho, notas por caso e sugestões de próximos passos. Esse relatório pode ser exportado em PDF, permitindo que o docente revise o percurso do estudante além da nota final do estudante. A interface contempla páginas de histórico, resultados e acompanhamento de tentativas, o que reforça o caráter de plataforma educacional.

Como evolução planejada, o EducaSim deverá incorporar novas teorias de aprendizagem, adaptações para perfis neurodivergentes e funcionalidades voltadas ao uso por professores, como criação de turmas, distribuição de avaliações e acompanhamento de desempenho por estudante. Essas extensões reforçam que o sistema foi pensado para apoiar formação profissional de modo progressivo e configurável.

5. Resultados e Discussões

Os resultados apresentados neste artigo referem-se à validação técnico-funcional do MVP, considerando o fluxo completo de criação da atividade, a sessão simulada, o relatório final e os testes do sistema. O primeiro eixo observado foi o sucesso da simulação de casos clínicos realistas. Em uma sessão completa na área de fisioterapia respiratória, o EducaSim conduziu cinco casos compatíveis com conteúdos curriculares de fisioterapia cardiopulmonar e pneumologia, incluindo DPOC, asma aguda, atelectasia em pós-operatório, bronquiolite e insuficiência respiratória com indicação de VNI. Os casos mobilizam sinais, parâmetros e condutas plausíveis, o que sugere aderência ao formato de capacitação pretendido e aproxima a experiência do contexto clínico real.

O segundo eixo foi a aderência ao material de referência informado. O EducaSim combina texto livre, páginas *web* e PDFs, processando essas fontes antes da avaliação. Embora esta versão do estudo não apresente uma métrica externa de fidelidade, a inspeção da transcrição da sessão, do relatório gerado e dos testes do sistema indica consistência temática entre referências, casos clínicos e devolutivas produzidas.

O terceiro eixo foi a correta aplicação da teoria de Kolb para guiar a experiência do usuário. A teoria não aparece como rótulo visível na conversa, mas estrutura cada caso em situação prática, reflexão, consolidação conceitual por *feedback* e nova aplicação. Esse padrão evita uma interação improvisada e transforma a teoria em mecanismo operacional, conectado a critérios de avanço, pistas, evidências esperadas e revisão da resposta.

O quarto eixo foi o funcionamento da avaliação interativa com rastreabilidade. O sistema acompanha o progresso do estudante, consolida pontuações por caso e gera relatório final com resumo do desempenho e próximos passos, deslocando a avaliação com IA de uma conversa efêmera para um processo documentado, passível de revisão docente. Esses resultados estabelecem a base técnico-funcional necessária para validações posteriores com especialistas e usuários, incluindo métricas de aprendizagem, comparação entre abordagens e análise da qualidade educacional das interações.

6. Considerações Finais

Este trabalho apresentou o EducaSim como uma plataforma *web* para avaliação interativa de casos clínicos com apoio de IA, concebida para reduzir limitações do uso manual de LLMs em atividades de capacitação. O diferencial da proposta está na integração entre ingestão multimodal de referências, planejamento pedagógico, geração validada de casos clínicos, interação guiada e geração de relatórios para acompanhamento docente. Em um primeiro momento, a plataforma implementa o ciclo de Kolb como teoria de aprendizagem configurável, organizando a experiência do estudante em torno de casos clínicos e reaplicação do conhecimento.

Os resultados preliminares indicam viabilidade técnico-funcional da solução, sobretudo quanto à simulação de casos, à aderência ao material de referência, à organização pedagógica da interação e à rastreabilidade do processo avaliativo. Ainda assim, o trabalho demanda etapas adicionais de teste e validação, especialmente com a participação de usuários, bem como investigações sobre ganhos de aprendizagem e ampliação do repertório pedagógico da plataforma.

Como próximos passos, pretende-se realizar avaliação por especialistas, conduzir

estudos com usuários, definir métricas de aprendizagem e ampliar o repertório de teorias de aprendizagem suportadas. Também estão previstas adaptações para perfis neurodivergentes e funcionalidades para docentes, como criação de turmas, compartilhamento de avaliações e acompanhamento sistemático do progresso dos estudantes.

Declaração sobre uso de IA generativa. Ferramentas de IA generativa foram utilizadas como apoio à revisão textual e à reorganização da redação deste manuscrito, sob supervisão dos autores.

Referências

- Abadam-Eremeev, D., Seng, R. P., Srinivasan, J., Knouna, K., Bogaudou, J. R., Rodriguez De La Nuez, P. A., and Podcheko, A. (2025). Revolutionizing educational assessment: The role of Bing Chat GPT-4 chatbot in increasing efficiency in grading open-ended questions. *Cureus*, 17(12):e100508.
- Asai, A., Wu, Z., Wang, Y., Sil, A., and Hajishirzi, H. (2023). Self-RAG: Learning to retrieve, generate, and critique through self-reflection. *arXiv preprint arXiv:2310.11511*.
- Aster, A., Laupichler, M. C., Rockwell-Kollmann, T., Masala, G., Bala, E., and Raupach, T. (2024). ChatGPT and other large language models in medical education – scoping literature review. *Medical Science Educator*.
- Flores, C. D., Caregnato, R. C. A., Wächter, D. A., Carvalho, P. F., Cogo, A. L. P., and Santos, E. F. d. (2023). Ferramenta de ensino para construção de casos clínicos do tipo paciente virtual. In *Anais Estendidos do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2023)*, pages 180–185, Porto Alegre. SBC.
- Guo, E., Ramchandani, R., Park, Y.-J., and Gupta, M. (2025). OSCEai: personalized interactive learning for undergraduate medical education. *Canadian Medical Education Journal*, 16(6).
- Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall, Englewood Cliffs, NJ.
- Loubbairi, S., El Moussaoui, Y., Lahlou, L., Chakri, I., and Nassik, H. (2025). The impact of artificial intelligence-driven simulation on the development of non-technical skills in medical education: a systematic review. *Journal of Educational Evaluation for Health Professions*, 22:37.
- Luo, M.-J., Pang, J., Bi, S., Lai, Y., Zhao, J., Shang, Y., et al. (2024). Development and evaluation of a retrieval-augmented large language model framework for ophthalmology. *JAMA Ophthalmology*, 142(9):798.
- Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegrefe, S., Alon, U., Dziri, N., Prabhunoye, S., Yang, Y., Gupta, S., Majumder, B. P., Hermann, K., Welleck, S., Yazdanbakhsh, A., and Clark, P. (2023). Self-Refine: Iterative refinement with self-feedback. *arXiv preprint arXiv:2303.17651*.
- Yu, H., Zhou, J., Li, L., Chen, S., Gallifant, J., Shi, A., et al. (2025). Simulated patient systems powered by large language model-based AI agents offer potential for transforming medical education. *Communications Medicine*, 6(1):27.