

ESTUDO COMPARATIVO DA PRECISÃO DA RESOLUÇÃO DE PROBLEMAS DE FUNÇÕES EM DIFERENTES MODELOS DE INTELIGÊNCIA ARTIFICIAL

Paloma Lohane de Melo Batista¹, Victor Valdez Fernandes¹, Estevão Vinícius Candia¹, Ewerton da Silva Schroeder¹

¹Instituto Federal de Mato Grosso do Sul – Campus Jardim - MS

paloma.batista@estudante.ifms.edu.br, victor.fernandes2@estudante.ifms.edu.br, estevao.candia@ifms.edu.br, ewerton.schroeder@ifms.edu.br

Área/Subárea: CET - Ciências Exatas e da Terra/Matemática

Tipo de Pesquisa: Científica

Palavras-chave: Inteligência Artificial. Resolução de problemas. Funções.

Introdução

A aplicação de Inteligências Artificiais (IA) no contexto educacional tem ganhado destaque como uma ferramenta potencial para aprimorar o aprendizado, especialmente em áreas desafiadoras como a matemática. No entanto, o uso inadequado dessa tecnologia, caracterizado pela busca de respostas rápidas sem o devido desenvolvimento do raciocínio lógico, levanta questionamentos sobre os efeitos a longo prazo na formação dos estudantes. Estudos como o de Camargo (2022), que analisou o impacto da IA no aprendizado de matemática em escolas públicas, e os de Cu e Hoffman (2023), que investigaram o uso de IA para resolver problemas matemáticos complexos, indicam uma crescente dependência dessas ferramentas. Além disso, Mariani et al. (2023) explora a influência da IA na redução da capacidade dos estudantes de aplicar raciocínios matemáticos em situações práticas.

O objetivo deste projeto é avaliar a precisão de modelos de Inteligência Artificial na resolução de problemas matemáticos na área de Funções, com foco em Polinomiais, Exponenciais e Logarítmicas, comparando-as com soluções humanas. A pesquisa busca entender o desempenho desses modelos no contexto educacional e suas possíveis implicações no processo de aprendizagem. Para isso, foram analisados diferentes modelos de IA e sua capacidade de resolver problemas de Funções, identificando suas limitações e acurácia. Assim, este estudo pretende fornecer uma visão crítica sobre o uso da IA como ferramenta pedagógica e seu impacto no desenvolvimento do raciocínio lógico dos alunos, apontando os riscos de dependência e a eficácia dessas tecnologias.

Metodologia

A pesquisa seguiu uma abordagem quantitativa, estruturada em etapas que garantiram uma análise precisa da acurácia de diferentes modelos de Inteligência Artificial (IA) na resolução de problemas matemáticos relacionados às Funções. Inicialmente, foi realizado um levantamento bibliográfico para a seleção criteriosa dos modelos de IA a serem testados, com foco naqueles capazes de resolver problemas de Funções Polinomiais, Exponenciais e Logarítmicas. Os modelos escolhidos para o estudo foram: Ayrton.ai, Bing, ChatGPT, Character.ai (Character Assistant e The Math Teacher), Gemini, LuzIA, Monica, Pi, You.com e ZapIA.

Para avaliar o desempenho dos modelos, foram extraídas 402 questões de livros didáticos, sendo 201 de Funções Polinomiais e 201 de Funções Exponenciais e Logarítmicas. As questões foram organizadas em dois formatos: abertas e de múltipla escolha, além de serem classificadas como com ou sem notação matemática. Essas mesmas questões foram resolvidas tanto pelos modelos de IA

quanto por humanos, utilizando as soluções dos livros como referência. A coleta de dados incluiu a análise das respostas fornecidas pelas IA e sua comparação com as soluções humanas. A precisão foi medida por meio de métricas como taxa de acerto e identificação de padrões de erros cometidos pelos modelos. O desempenho dos modelos foi registrado e os dados tabelados foram processados estatisticamente para quantificar sua acurácia e identificar suas limitações.

Além da análise quantitativa, foram realizadas análises qualitativas para identificar dificuldades enfrentadas pelos modelos e padrões de comportamento na resolução dos problemas. Todos os resultados foram organizados em um relatório final, que destaca o impacto do uso de IA no ensino de matemática e sugere direções para pesquisas futuras.

Resultados e Análise

Alguns dos modelos de Inteligência Artificial analisados, como o ChatGPT, Bing, Monica e LuzIA, destacaram-se por apresentar um desempenho relativamente superior em comparação aos outros, especialmente na resolução de problemas de Funções Polinomiais. O ChatGPT, por exemplo, alcançou uma taxa de acerto de 68,66% nesse tipo de função, o que o posiciona como o modelo com melhor performance nessa categoria. No entanto, ao se analisar o desempenho em Funções Exponenciais e Logarítmicas, o cenário se torna mais preocupante: o Bing, que obteve a maior taxa de acerto nessa área, registrou apenas 48,26%. Esses números indicam que, embora alguns modelos tenham se saído melhor em certos tipos de problemas, o desempenho geral ainda está longe de ser satisfatório. A Tabela 1 apresenta o ranking dos modelos com base nas suas taxas de acerto, e mostra que nenhum deles atingiu níveis de precisão que possam ser considerados confiáveis o suficiente para uso em um contexto educacional sem supervisão. A variabilidade de desempenho entre os diferentes tipos de funções e a presença constante de erros, mesmo em questões simples, levantam questionamentos sobre a viabilidade desses modelos para auxiliar efetivamente no aprendizado de matemática.

Além disso, o estudo revelou uma questão crucial: a frequência de respostas absurdas, que variaram desde erros conceituais a incoerências matemáticas. Esses erros não se limitaram a problemas complexos; em várias ocasiões, os modelos falharam em resolver corretamente problemas relativamente simples, sugerindo uma limitação clara em sua capacidade de interpretar e aplicar os conceitos matemáticos fundamentais. A seguir, serão apresentados dois exemplos de respostas absurdas fornecidas pelos modelos, que evidenciam essa falta de confiabilidade.

Funções Polinomiais			Funções Exponenciais e Logarítmicas		
POS.	IA	%	POS.	IA	%
1°	ChatGPT	68,66	1°	Bing	48,26
2°	Monica	62,69	2°	ChatGPT	47,76
3°	Bing	58,71	3°	LuzIA	47,26
4°	LuzIA	54,73	4°	Monica	45,77
5°	Gemini	49,75	5°	You.com	40,80
6°	You.com	45,77	6°	Gemini	35,32
7°	ZapIA	40,30	7°	ZapIA	26,87
8°	Ayrton.ai	24,88	8°	Ayrton.ai	25,37
9°	Character.ai - Character Assistant	15,92	9°	Pi	24,88
10°	Character.ai - The Math Teacher	6,97	10°	Character.ai - Character Assistant	10,95
11°	Pi	3,48	11°	Character.ai - The Math Teacher	6,47

Tabela 1. Ranking do desempenho das Ias pela porcentagem de acertos

A questão FP0018, por exemplo, perguntava qual era a medida de comprimento de uma coluna de mercúrio em termômetro quando a medida de temperatura corpórea do paciente era de 39 °C, dadas algumas informações. A resposta correta é 7,9 cm. No entanto, a resposta fornecida pela IA Ayrton.ai foi -0,1 cm, o que não faz sentido algum.

Outro exemplo é a questão EL0176, que a partir de cálculos logarítmicos gostaria de saber em que ano a população da Terra ultrapassaria 7 bilhões de habitantes. A resposta de referência era no decorrer de 2009. No entanto, a IA Bing forneceu como resposta 2034,2.

Esses exemplos demonstram que, embora a IA tenha o potencial de auxiliar no ensino de Matemática, sua precisão e consistência ainda são insuficientes para garantir uma aplicação efetiva e segura no ambiente educacional sem o risco de prejudicar o processo de aprendizagem dos estudantes. Assim, os resultados desta pesquisa ressaltam a importância de uma análise mais detalhada sobre o desempenho dos modelos de IA em diferentes contextos matemáticos, e apontam para a necessidade de um desenvolvimento mais robusto dessas ferramentas antes que possam ser plenamente integradas ao sistema educacional.

Considerações Finais

Embora os modelos de IA tenham demonstrado algum nível de eficácia, principalmente em Funções Polinomiais, eles ainda não são suficientemente precisos para substituir o ensino tradicional. Os alunos devem possuir um conhecimento sólido dos conceitos matemáticos para identificar e corrigir respostas incorretas fornecidas pela IA. A dependência excessiva dessas ferramentas pode comprometer o aprendizado. Portanto, seu uso deve ser complementar e supervisionado por professores, garantindo que o foco continue no desenvolvimento do raciocínio lógico dos estudantes.

Agradecimentos

Agradecemos ao nosso orientador, Estevão Vinícius Candia, e ao coorientador, Ewerton da Silva Schroeder, por sua orientação e apoio fundamentais ao longo deste projeto. Também expressamos nossa gratidão ao IFMS - Campus Jardim pelo incentivo à pesquisa. Por fim, agradecemos ao CNPq pelo apoio financeiro através da bolsa de pesquisa, que contribuiu para a realização deste trabalho.

Referências

CAMARGO, A. **Professores, atenção! ChatGPT é usado como cola nas escolas.** 2022. Disponível em: <https://is.gd/NZw4Y3>. Acesso em: 13 set. 2024.

CU, M. A.; HOCHMAN, S. **Scores of Stanford students used ChatGPT on final exams, survey suggests.** The Standford Daily: Science & Technology. Standford, p. 1-3. jan. 2023.

MARIANI, D.; YUKARI, D.; DIEGUES, L.; SOPRANA, P.; HERNANDES, R.; **ChatGPT vai melhor que 80% dos alunos no Enem, mas derrapa em matemática.** Folha de São Paulo [online], São Paulo, 05 abr.2023. Tecnologia. Disponível em: <https://is.gd/I0ti1R>. Acesso em: 13 set. 2024.

COMPARATIVE STUDY OF THE ACCURACY OF PROBLEM SOLVING IN FUNCTIONS ACROSS DIFFERENT ARTIFICIAL INTELLIGENCE MODELS

Abstract: Artificial Intelligence (AI) models have gained popularity in various fields. Students have been using these models as a way to circumvent the creative and evaluative processes in multiple areas of knowledge. This research project aims to investigate the accuracy of AIs in solving problems related to Functions. A quantitative approach will be adopted, collecting data through the selection of AI models and mathematical problems for analysis. It is expected that this research will provide data demonstrating the effectiveness (or lack thereof) of these AI models in this area and establish a foundation for future studies, in addition to improving their applicability and reducing their limitations.

Keywords: Artificial Intelligence. Problem Solving. Functions.