

15º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2024

Analisando o ChatGPT-3.5 como ferramenta para resolução de exercícios de física.

ALLAN FERRAZ¹, ALEX LINO²

¹ Graduando em Licenciatura em Física, Bolsista PIBIT, IFSP, Campus Caraguatatuba, Allan.Ferraz@aluno.ifsp.edu.br.

² Doutor em Educação para a Ciência pela UEM. Professor do Instituto Federal de São Paulo – IFSP, campus Caraguatatuba, Alex.lino@ifsp.edu.br

Área de conhecimento (Tabela CNPq): 7.08.04.00-1 Ensino-Aprendizagem

RESUMO: O presente trabalho teve como foco a avaliação da viabilidade do ChatGPT-3.5 na resolução de problemas e exercícios de Física. A confiabilidade do uso do ChatGPT-3.5 como ferramenta na Física foi examinada por meio da construção de um banco de dados contendo as questões teóricas e algébricas. Foram analisados os percentuais de acerto e a coerência das respostas de forma qualitativa, explorando possíveis causas para os erros identificados. Concluindo que o uso dessa inteligência artificial requer um conhecimento prévio em Física por parte dos usuários que a utilizam para resolução de problemas em Física.

PALAVRAS-CHAVE: ChatGPT-3.5. Ensino de Física. Tecnologia da Informação e Comunicação.

Analyzing ChatGPT-3.5 as a tool for solving physics exercises.

ABSTRACT: This work focused on evaluating the feasibility of ChatGPT-3.5 in solving physics problems and exercises. The reliability of using ChatGPT-3.5 as a tool in physics was examined by constructing a database containing theoretical and algebraic questions. The accuracy rates and the coherence of the responses were analyzed qualitatively, exploring possible causes for the identified errors. Concluding that the use of this educational tool requires prior knowledge in physics.

KEYWORDS: ChatGPT-3.5. Physics Teaching. Informations and Communications Technology.

INTRODUÇÃO

Fatores como as práticas que incentivam a memorização e a aprendizagem mecânica, a repetição excessiva de lições em sala de aula, a descontextualização e a falta de problematização dos conteúdos científicos, bem como o foco na resolução de exercícios com ênfase exagerada na aplicação de "fórmulas", entre outros aspectos, contribuem para a crise no Ensino de Física e no Ensino de Ciências em geral (Bertusso et al., 2020).

De modo geral a crise no Ensino de Física e no Ensino de Ciências em geral tem sido um tema que preocupa pesquisadores da área, e muito trabalho vem sendo realizado e divulgado, desde os anos 1980, com objetivo de buscar soluções a esse problema (MOREIRA, 2018).

Devemos levar em consideração que as diversas tecnologias contemporâneas estão atingindo cada vez mais intensamente as pessoas em geral, em diferentes camadas da sociedade, e as TICs, acompanhando esse movimento, impactam de forma significativa trazendo significativas transformações na vida das pessoas. No contexto educacional não é diferente. Já é realidade o que o ambiente escolar enfrenta sérios desafios com a utilização dessa tecnologia por parte dos professores e estudantes (Dourado et al., 2015). Esse desafio converge basicamente ao modo como essa tecnologia tem sido utilizada fora do ambiente escolar e, tendo uma sociedade cada vez mais conectada, devemos refletir sobre esse novo processo de ensino-aprendizagem.

Moreira (2018) defende que as TICs devem estar mais presentes nesse processo de ensino e substituir os modelos tradicionais de ensino que envolvem excesso de utilização de fórmulas,

matematização, processos de aprendizagem mecânica que não levam em consideração os erros dos estudantes.

TICs ou Tecnologias da informação e comunicação, são ferramentas e recursos tecnológicos utilizados para o processamento, armazenamento e disseminação de informações. Incluem desde dispositivos físicos, como computadores e smartphones, até softwares, redes e serviços digitais que facilitam a comunicação e a gestão de dados, até as mais novas inteligências artificiais ou IA como o ChatGPT. As TICs desempenham um papel crucial na modernização de processos, na automação de tarefas e na conectividade global, impactando diversos setores, como a educação.

Com o avanço das TICs e das Inteligências Artificiais, o ChatGPT se destacou por ter sido utilizado amplamente no contexto educacional, tanto por parte dos professores como pelos estudantes. De forma geral os estudantes podem utilizar a ferramenta para resolução de exercícios, definições de conceitos ou para confecção de trabalhos ou relatórios. Por esse motivo se faz necessário a consciência por parte dos professores sobre os modos de utilização da ferramenta e os possíveis obstáculos e riscos que ela pode oferecer ao aprendizado dos estudantes. Diante desse cenário, o principal objetivo desse trabalho foi o de verificar a viabilidade dessa IA como ferramenta educacional para resolução de exercícios no campo da Física.

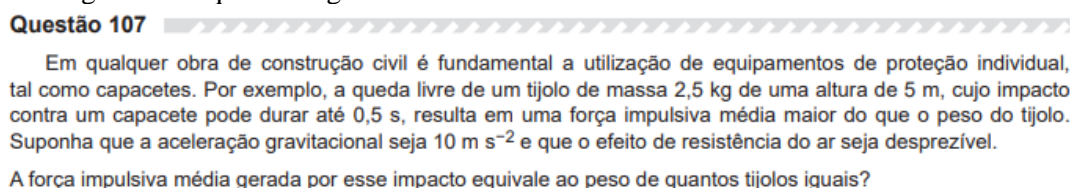
MATERIAL E MÉTODOS

O método empregado para analisar o desempenho do ChatGPT-3.5 foi dividido em duas etapas principais. A primeira etapa envolveu a criação de um banco de dados contendo questões de física, metade sendo questões teóricas e metade sendo questões algébricas.

A segunda etapa consistiu em uma análise qualitativa das respostas fornecidas pela IA permitindo assim a obtenção das taxas de acerto do ChatGPT-3.5 em relação as questões propostas.

Na fase inicial de coleta de dados, foram escolhidas 40 questões de Física oriundas de diversas fontes, como provas anteriores do ENEM (Exemplo na Figura 1) e sites educacionais, incluindo "TodaMatéria.com" e "BrasilEscola.com". Essas questões foram organizadas em duas categorias: vinte questões conceituais, que não demandavam operações algébricas para sua resolução, e vinte questões algébricas, que exigiam cálculos matemáticos. Posteriormente, solicitamos ao ChatGPT-3.5 que gerasse 10 respostas para cada uma das 40 questões selecionadas (Exemplo na Figura 2). Os critérios de seleção das questões foram definidos pela ausência de necessidade de interpretação de imagens ou gráficos, pela natureza física das questões e pela relevância do conteúdo para o ensino médio. Com isso, foram gerados 400 dados para análise, resultantes das 10 respostas do ChatGPT-3.5 para cada uma das 40 questões.

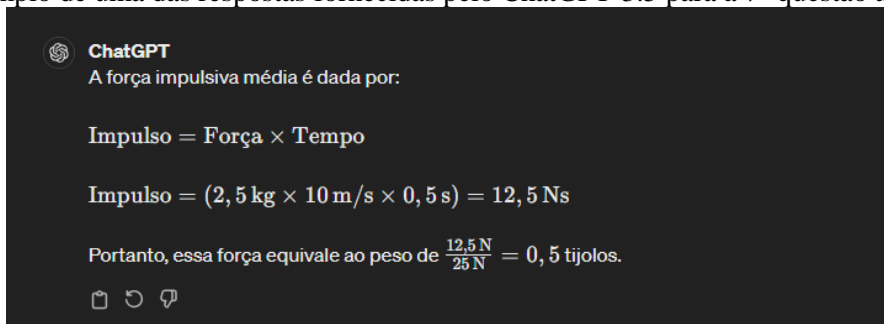
FIGURA 1. Imagem da 7ª questão algébrica.



Fonte: ENEM (2019, Caderno amarelo)

Na Figura 1, podemos ver uma das 40 questões selecionadas para o uso no banco de dados. As questões tiveram suas alternativas de respostas retiradas, para que não influenciassem as respostas do ChatGPT-3.5. E após isso, sem nenhuma alteração no enunciado, as questões foram submetidas ao ChatGPT-3.5.

FIGURA 2. Exemplo de uma das respostas fornecidas pelo ChatGPT-3.5 para a 7ª questão algébrica.



Fonte: ChatGPT-3.5 (2023)

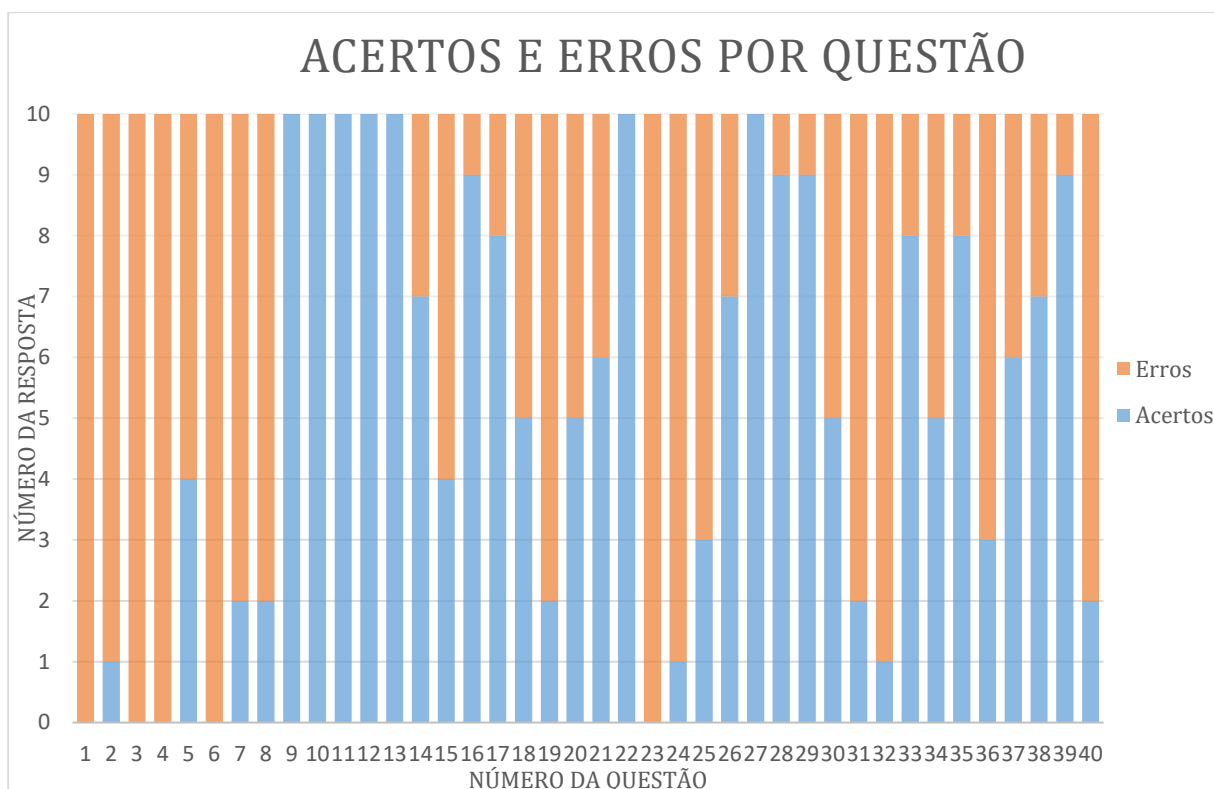
Na figura 2, vemos uma das dez respostas do ChatGPT-3.5 para a questão da figura 1, cada resposta fornecida por ele, foi avaliada, se estava correta ou não, e caso estivesse errada, foi avaliada onde ocorreu esse erro, e do por que o erro foi cometido. E após esse processo, cada resposta foi categorizada e adicionada no banco de dados criado.

RESULTADOS E DISCUSSÃO

A análise inicial considerou os 400 dados obtidos, compostos por questões conceituais e algébricas.

Dos 400 itens, o ChatGPT-3.5 acertou 210 e errou 190. Observou-se que a maioria dos acertos ocorreu nas questões conceituais. Para ilustrar essa distribuição, foi elaborado o Gráfico 1.

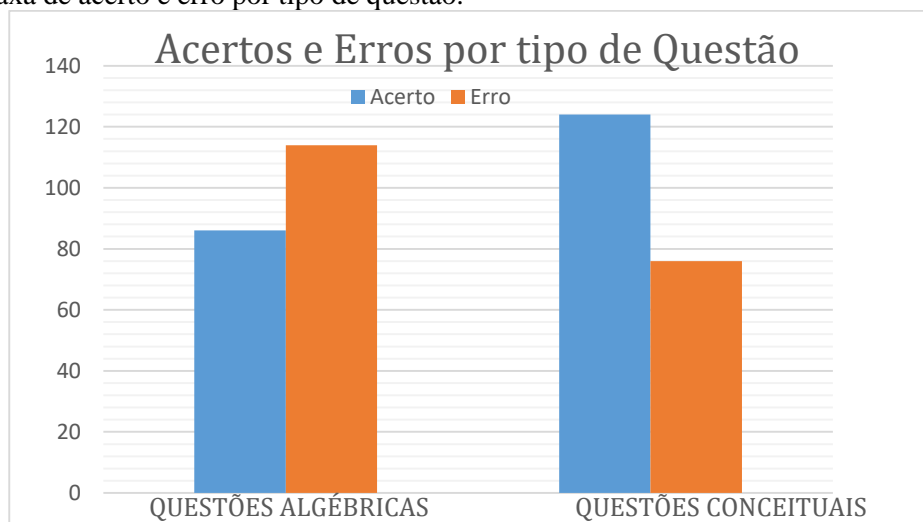
Gráfico 1: Taxa de acerto e erro por questão.



Fonte: Autor.

Ao examinar o Gráfico 1, observa-se que, das 400 questões analisadas, o ChatGPT obteve 210 respostas corretas e 190 incorretas. Notamos também que a maioria dos acertos ocorreu em questões de natureza conceitual. Com base nisso, foi elaborado o Gráfico 2, que categoriza as questões de acordo com os dois tipos específicos previamente selecionados.

Gráfico 2: Taxa de acerto e erro por tipo de questão.



Fonte: Autor.

O Gráfico 2 oferece uma análise detalhada, demonstrando que o ChatGPT-3.5 alcançou uma taxa de acerto mais elevada em questões de natureza conceitual (62%), em contraste com as questões que exigiam habilidades de algebrismo, onde a taxa foi de 43%. A avaliação dos dados revela uma precisão geral de 52% (210 respostas corretas), contra 48% de erros (180 respostas). Esta análise preliminar sugere que, embora o ChatGPT-3.5 tenha gerado mais respostas corretas do que incorretas, sua eficiência como ferramenta de soluções de questões de Física, pode ser limitada.

Análise dos erros:

Durante o processo de resposta do ChatGPT-3.5, foram identificados diversos erros, tanto em aspectos algébricos quanto interpretativos.

De acordo com Diogo Cortiz (2023), um dos erros mais recorrentes cometidos pelo ChatGPT-3.5 são as chamadas "alucinações". Esse termo se refere a situações em que o modelo inclui informações ou dados inexistentes no enunciado original, resultando em respostas incorretas ou inviáveis para a resolução do problema. Esse fenômeno é particularmente notável em questões algébricas, onde valores fictícios foram inseridos no processo de resolução, causando confusão e inviabilizando a solução correta.

Por exemplo, em uma questão algébrica que solicitava a resolução de um problema simples sobre um elevador hidráulico com pistões de tamanhos diferentes, o problema poderia ser resolvido corretamente utilizando o teorema de Pascal. No entanto, em algumas respostas, o ChatGPT-3.5 apresentou alucinações ao introduzir fatores como o arrasto do líquido na tubulação, o que resultou em respostas incoerentes e afastadas da solução correta.

Outro tipo de erro comum observado foi a "Confusão Generalizada". Esse erro ocorreu frequentemente em questões que exigiam múltiplos passos para a resolução ou envolviam a conversão de unidades. Na tentativa de simplificar o processo, o ChatGPT-3.5 frequentemente aplicou os valores fornecidos de maneira incorreta, comprometendo a precisão dos resultados.

Esse fenômeno ficou evidente em várias respostas onde a solução correta dependia da integração de várias etapas ou da manipulação adequada de unidades de medida. Ao lidar com problemas complexos, como aqueles que exigem a aplicação de fórmulas com múltiplas variáveis ou a conversão entre diferentes unidades, o ChatGPT-3.5 por vezes não conseguiu manter a coerência entre os diferentes passos necessários. Em casos envolvendo a aplicação de fórmulas físicas ou conversão de unidades, o modelo pode ter misturado unidades, aplicado fórmulas de forma inadequada ou desconsiderado elementos cruciais do enunciado, levando a respostas incorretas ao simplificar indevidamente o processo, sem respeitar a complexidade e precisão exigidas para a resolução adequada dos problemas.

CONCLUSÕES

A proposta deste estudo foi avaliar a eficácia do ChatGPT-3.5 como ferramenta para a solução de questões de Física. É fundamental enfatizar que os resultados e análises apresentados são específicos para o ChatGPT-3.5 e não podem ser extrapolados para outras versões do modelo, como o ChatGPT-4.0 ou futuras atualizações.

A partir da análise dos dados e dos erros encontrados, ficou claro que o ChatGPT-3.5 enfrenta desafios consideráveis ao lidar com questões de Física, sejam elas algébricas ou teóricas. É importante observar que o ChatGPT-3.5, obteve uma taxa de precisão de apenas 52% nas questões algébricas e conceituais, com uma diferença considerável de acerto das questões algébricas (43%) comparado as conceituais (62%). Demonstrando assim uma maior facilidade da ferramenta ao trabalhar com a parte teórica da Física, mas não o suficiente para realmente ser uma ferramenta confiável para a área de resoluções de exercícios de Física.

Considerando os resultados deste estudo, e os erros observados na análise (sejam eles dado por alucinações do ChatGPT-3.5 ou por confusões criadas pelo mesmo) a recomendação para o uso do ChatGPT-3.5 como ferramenta para soluções de exercícios de Física enfrenta problemas notáveis, principalmente devido à sua tendência de gerar erros que podem levar a confusão. No entanto, se utilizado por uma pessoa experiente na área, que tenha a capacidade de detectar e corrigir essas falhas, o ChatGPT-3.5 pode se mostrar uma ferramenta útil. Podendo criar diversas soluções diferentes para o mesmo problema em apenas alguns instantes. A eficácia do ChatGPT-3.5 é, portanto, amplamente determinada pela criatividade do usuário. Apesar disso, é crucial usar a ferramenta com cautela e sempre validar as informações fornecidas para garantir a precisão.

CONTRIBUIÇÕES DOS AUTORES

Autor 1 realizou o trabalho sob orientação do autor 2. Ambos os autores contribuíram com a escrita e revisão deste trabalho, aprovando a versão submetida.

AGRADECIMENTOS

Agradecemos ao financiamento e apoio do PIBITI e ao IFSP – Campus Caraguatatuba.

REFERÊNCIAS

BERTUSSO, F. R. et al. A utilização das Tecnologias de Informação e Comunicação (TICs) no ensino de Ciências: um paradigma a ser vencido. **Research, Society and Development**, Vargem Grande Paulista, v. 9, n. 12, p. 1-18, dez. 2020. Disponível em: https://www.researchgate.net/publication/347943445_A_utilizacao_das_Tecnologias_de_Informacao_e_Comunicacao_TICs_no_ensino_de_Ciencias_um_paradigma_a_ser_vencido. Acesso em: 01 jul. 2024.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. (BNCC) Brasília, 2018.

CORTIZ, Diogo. **Entendendo as Alucinações do ChatGPT**. Diogo Cortiz, 2023. Disponível em: <https://diogocortiz.com.br/entendendo-as-alucinacoes-do-chatgpt/>. Acesso em 01 jun. 2024.

DOURADO, Irismar de França; DE SOUZA, Keith Leandro; CARBO, Leandro; MELLO, Geison Jader; AZEVEDO, Lucy Ferreira. Uso das TIC no Ensino de Ciências na Educação Básica: uma Experiência Didática. **Revista de Ensino, Educação e Ciências Humanas**, Londrina, v. 15, 2015. DOI: 10.17921/2447-8733.2014v15n0p%p. Disponível em: <https://revistaensinoeducacao.pgsscogna.com.br/ensino/article/view/438>. Acesso em: 28 ago. 2024.

MOREIRA, Marco Antônio. Uma análise crítica do ensino de Física. **Estudos avançados**, São Paulo, v. 32, n. 94, p. 73-80, 2018. Disponível em: <https://www.scielo.br/j/ea/a/3JTLwqQNsfWPqr6hjzyLQzs>. Acesso em 1 jun. 2024.