

## INTELIGÊNCIA ARTIFICIAL APLICADA NA CORREÇÃO DE AVALIAÇÕES: LIMITES, LIÇÕES E POSSIBILIDADES

Tiago Brandão Costa, Universidade Federal Fluminense - UFF, [tiagobrandao@id.uff.br](mailto:tiagobrandao@id.uff.br),  
Tatiana Caneda Salazar, Universidade Federal Fluminense - UFF, [tatianasalazar@id.uff.br](mailto:tatianasalazar@id.uff.br),  
Denise Hirayama, Universidade Federal Fluminense - UFF, [denisehirayama@id.uff.br](mailto:denisehirayama@id.uff.br),  
Ésoly Madeleine Bento dos Santos, Universidade Federal Fluminense - UFF,  
[esolysantos@id.uff.br](mailto:esolysantos@id.uff.br)

Andrei Bonamigo, Universidade Federal de Santa Catarina - UFSC, [andreibonamigo@gmail.com](mailto:andreibonamigo@gmail.com)

### Resumo

Esse relato de experiência apresenta a avaliação do ChatGPT como ferramenta auxiliar na correção de avaliações das disciplinas de Estrutura de materiais e Propriedades de materiais do ciclo básico do curso de Engenharia da UFF, Campus Volta Redonda - RJ. Diante da sobrecarga de trabalho docente e seus impactos na saúde mental e melhoria contínua do processo, o experimento buscou avaliar a viabilidade da inteligência artificial (IA) na correção de provas manuscritas. Foram testadas diferentes estratégias, incluindo o uso de arquivos em JPEG, *prompts* e *chat* temporário. Os resultados mostraram que a IA teve bom desempenho em questões objetivas simples, mas apresentou limitações na leitura de manuscritos, análise de gráficos e correção de questões complexas. Também foram observadas “alucinações” e notas inconsistentes. Com base nos resultados, pode-se observar que a IA pode ser uma aliada no processo avaliativo, mas não substitui o julgamento humano qualificado, sendo necessária mediação docente e formação adequada para seu uso responsável.

**Palavras-chave:** ChatGPT, IA, Engenharia, Correção de avaliação.

## 1. Introdução

O desgaste psicológico entre professores tem sido amplamente observado nos últimos anos, especialmente no ensino superior (Santos; Cortez, 2021; Lima; Paloski, 2024; Wagner et al., 2021; Moura et al., 2019). Estudos indicam que a sobrecarga de trabalho está relacionada à multiplicidade de funções desempenhadas pelos docentes para além das atividades em sala de aula, como participação em eventos acadêmicos, reuniões institucionais, gestão de cursos, planejamento didático, correção de avaliações e produção de documentos (Moura et al., 2019). Essa rotina intensa é agravada por fatores como a necessidade de atenção constante, o excesso de responsabilidades, a rivalidade entre colegas, a perda de autonomia, o cansaço físico e mental, cobranças frequentes e a remuneração insatisfatória (Wagner et al., 2021).

Consequentemente, pesquisas têm apontado altos índices de adoecimento psicológico entre professores, com destaque para sintomas de depressão, ansiedade e outras manifestações emocionais que afetam diretamente sua saúde e desempenho profissional. Relatos de afastamentos por motivos de saúde, desmotivação e insatisfação no trabalho são recorrentes (Wagner et al., 2021). Entre os principais fatores apontados pelos docentes, destaca-se o acúmulo de funções, particularmente no ensino superior, onde há a exigência de constante acompanhamento dos estudantes e a pressão por produtividade acadêmica, incluindo a elaboração e publicação de pesquisas (Santos; Cortez, 2021).

Dados recentes reforçam a gravidade desse cenário. Lima e Paloski (2024) verificaram que, entre os docentes de instituições públicas, 26,68% apresentavam sintomas de depressão, 17,34% de ansiedade e 28,67% de estresse, variando entre níveis leves, moderados, severos ou extremamente severos. Entre os docentes de instituições privadas, os percentuais foram ainda mais elevados: 32,73%, 20,01% e 20,99%, respectivamente. Esses resultados evidenciam a alta prevalência de transtornos mentais entre professores universitários e ressaltam a importância de ações voltadas ao cuidado com a saúde mental docente. A similaridade nos estilos de ensino entre instituições públicas e privadas reforça a necessidade urgente de suporte psicológico e educacional abrangente, promovendo o bem-estar no ambiente acadêmico (Lima; Paloski, 2024).

Além dos dados quantitativos que revelam a alta incidência de sofrimento psíquico entre docentes, destaca-se também a complexidade dos fatores envolvidos nesse adoecimento. O ambiente escolar, em todos os níveis de ensino, tem se revelado um espaço de constantes pressões, marcado por desvalorização profissional, pouco reconhecimento institucional, conflitos nas relações interpessoais, ausência de apoio familiar aos alunos e precariedades físicas nas condições de

trabalho. Tais elementos contribuem diretamente para o agravamento de quadros como a síndrome do pânico, *Burnout* e outras manifestações emocionais. Em muitos casos, os professores enfrentam jornadas ampliadas, ausência de momentos de lazer e sentimento de frustração por não conseguirem conciliar as demandas profissionais com a vida pessoal. Esse cenário, para além de comprometer o bem-estar dos docentes, interfere na qualidade do ensino e fragiliza o vínculo estabelecido com os estudantes. Diante disso, torna-se essencial implementar políticas de cuidado, oferecer suporte emocional contínuo e criar espaços de escuta e acolhimento nas instituições educacionais. Valorizar o professor, em sua totalidade humana, é também investir na qualidade da educação (Carvalho, Lins e Gusmão, 2025).

Nesse contexto, o uso de ferramentas baseadas em inteligência artificial (IA) no campo educacional surge como uma possível aliada para minimizar alguns desses desgastes. A IA promete revolucionar a forma como os professores lidam com correções, avaliações e *feedbacks*. Sistemas automatizados já demonstram eficiência na correção de provas e trabalhos, agilizando processos repetitivos e liberando tempo para que os educadores se dediquem ao que realmente importa: o ensino e o acompanhamento personalizado dos alunos. Ferramentas de IA podem identificar padrões de erro em turmas inteiras, emitir *feedback* instantâneo e até auxiliar na gestão administrativa, reduzindo significativamente a sobrecarga de tarefas burocráticas (Okagbue et al., 2023).

No entanto, a adoção dessas tecnologias não é isenta de desafios. Embora facilitem a correção em larga escala, sua precisão ainda depende de um treinamento rigoroso e da supervisão humana para evitar vieses e interpretações equivocadas. Um exemplo é o estudo de Santos (2023), que avaliou o desempenho do ChatGPT-3.5 na resolução de questões de Matemática do ENEM entre 2017 e 2022. Os resultados revelaram limitações significativas do modelo nessa área, com uma taxa média de acerto de apenas 45,05% com alternativas, e queda para 31,57% em questões abertas.

As principais deficiências observadas foram: dificuldade em interpretar problemas complexos (especialmente em geometria descritiva), falhas na construção de raciocínios lógicos e erros básicos de cálculo. Isso evidencia que, embora o ChatGPT seja capaz de gerar respostas coerentes em linguagem natural, ainda não compreende verdadeiramente os conceitos matemáticos, reproduzindo padrões aprendidos em seu treinamento. Comparado a outros modelos, como o Bard, modelo baseado na família de modelos de linguagem LaMDA (*Language Model for Dialogue Application*), os resultados também foram insatisfatórios, reafirmando que, em questões que exigem precisão e pensamento estruturado, a avaliação humana ainda é indispensável (Santos, 2023). Esse fato corrobora com os resultados obtidos por Celestino e Valente (2024). De acordo com os autores, a

principal dificuldade do ChatGPT, está relacionada às suas limitações em matemática e análise estatística. Embora a ferramenta seja útil para fornecer explicações passo a passo sobre o uso de softwares e análise de dados, ela foi projetada com foco no uso linguístico. Isso significa que suas respostas podem apresentar erros, inconsistências, perda de contexto e viés, especialmente em conteúdos quantitativos mais complexos (Celestino e Valente, 2024).

Em contrapartida, o uso do ChatGPT na resolução da prova da Ordem dos Advogados do Brasil (OAB) demonstrou que com o treinamento constante pode-se aprimorar sua capacidade de análise e precisão. Após ser exposto a milhares de questões do Exame da Ordem, o modelo passou a reconhecer padrões, interpretar enunciados complexos e oferecer respostas mais alinhadas ao esperado pela banca. A prática repetitiva mostrou-se essencial para que a IA refinasse sua compreensão, corrigisse erros e se tornasse mais eficaz um processo que reflete o princípio de que a prática leva à perfeição, também válido para os humanos (Inteligência artificial é aprovada em 1ª fase da OAB, 2024).

O ChatGPT, é uma Inteligência Artificial desenvolvida pela OpenAI e lançado em 30 de novembro de 2021, é um chatbot baseado em modelos de linguagem de larga escala (*LLMs – Large Language Models*) da família GPT (*Generative Pre-trained Transformer*). Treinado em uma vasta base de dados textual, o ChatGPT é capaz de interagir de forma contextualizada e gerar respostas coerentes em diversos domínios. A ferramenta tem demonstrado bom desempenho em tarefas que exigem compreensão textual, produção escrita, interpretação de questões e resolução de problemas inclusive, superando em muitos casos a média humana em exames padronizados (OPENAI, 2023).

Diante desse cenário, a prática aqui relatada consistiu na aplicação da IA ChatGPT 4.0 (versão gratuita) para corrigir provas já avaliadas previamente pelos professores. O objetivo desta experiência foi avaliar a viabilidade do uso do ChatGPT – 4.0 como ferramenta auxiliar na correção de avaliações, considerando sua capacidade de interpretar respostas manuscritas (digitalizadas) e atribuir notas coerentes com os critérios definidos pelo docente. Além disso, buscou-se comparar os resultados obtidos pela IA com as notas previamente atribuídas pelo professor, a fim de verificar a precisão, a confiabilidade e as limitações do uso da IA nesse contexto.

A seguir, este relato apresenta os resultados da aplicação da metodologia descrita, discutindo suas implicações, benefícios e eventuais obstáculos observados no uso da inteligência artificial como apoio à prática docente na educação superior.

## 2. Descrição do problema

O presente relato de experiência objetiva analisar a viabilidade do uso da ferramenta de inteligência artificial ChatGPT na correção de avaliações aplicadas a discentes dos cursos de Engenharia da Escola de Engenharia Industrial Metalúrgica de Volta Redonda (EEIMVR) da Universidade Federal Fluminense (UFF). A motivação surgiu da necessidade de aperfeiçoar o tempo e o esforço demandados pelo processo de correção de provas, principalmente em disciplinas com alto número de estudantes e questões de natureza objetiva e discursiva. Identificou-se como desafio central a capacidade da ferramenta de interpretar diferentes tipos de questões objetivas e discursivas bem como avaliar corretamente provas manuscritas, algumas contendo gráficos, figuras e desenhos.

## 3. Solução desenvolvida (percurso metodológico)

Para investigar as potencialidades e limitações do uso da inteligência artificial como apoio na correção de avaliações em cursos de Engenharia, foi conduzido um estudo aplicado com base em provas reais de duas disciplinas de natureza teórica, pertencentes à matriz curricular dos cursos de Engenharia. As avaliações haviam sido respondidas de forma manuscrita pelos alunos, em folhas A4, e continham questões diversificadas tanto no formato quanto no conteúdo das questões. Foram avaliadas: 7 questões múltipla escolha (5 apenas textuais e 2 com análise gráfica ou de imagens), 1 questão discursivas com interpretação de gráficos e/ou tabelas, 2 outras discursivas exigindo a elaboração de desenhos técnicos e por fim 2 questões discursivas exclusivamente textuais. As Figuras de 1 a 4 ilustram essa variedade.

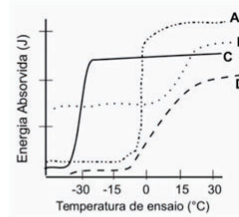
Figura 1 - Questão de múltipla escolha textual

- 1) (Valor 0,8 pontos)** A recristalização de um material deformado depende do tamanho de grão inicial, da intensidade de deformação, da temperatura de deformação e da temperatura de recristalização. Nesse processo de recristalização, ao se
- (A) reduzir a intensidade de deformação, reduz-se a temperatura de recristalização para um mesmo tempo de tratamento.
  - (B) reduzir a intensidade de deformação, diminui-se o tamanho do grão recristalizado para uma mesma temperatura e tempo de tratamento.
  - (C) aumentar a intensidade de deformação, diminui-se o tamanho do grão recristalizado para uma mesma temperatura e tempo de tratamento.
  - (D) aumentar a intensidade de deformação, aumenta-se a temperatura de recristalização para um mesmo tempo de tratamento.
  - (E) aumentar a intensidade de deformação, aumenta-se o tamanho do grão recristalizado para uma mesma temperatura e tempo de tratamento.

Fonte: os autores

Figura 2 - Questão de múltipla escolha com imagem

2) (Valor 0,8 pontos) A Figura ao lado apresenta quatro curvas experimentais de energia absorvida em função da temperatura de Ensaio Charpy para quatro aços distintos. Um engenheiro precisa escolher dois aços para serem empregados na construção de estruturas que podem sofrer impactos. A construção terá lugar numa região cuja menor temperatura ambiente registrada nos últimos 100 anos foi 0°C, e cuja temperatura média é de 20°C. Os aços que devem ser escolhidos pelo engenheiro, considerando somente o ensaio Charpy, são aços



- (A) A e B
- (B) B e C
- (C) A e C
- (D) B e D
- (E) C e D

Fonte: os autores

Figura 3 - Questão discursiva com solicitação de desenho técnico

6) (Valor 4,0 pontos) A escolha do material ideal depende dos requisitos de desempenho, custo e durabilidade da aplicação. É fundamental **compreender as características e limitações de cada tipo de material para otimizar o desenvolvimento de produtos e soluções inovadoras**. A partir do seu conhecimento sobre materiais cerâmicos, metálicos e poliméricos faça o que se pede nas seguintes questões:

- a) (1,0 ponto) desenhe uma curva padrão de tensão-deformação de um ensaio de tração uniaxial para cada grupo de materiais (cerâmico, metálico e polímero semicristalino) e identifique nessa curva, com as letras a, b, c e d os seguintes pontos notáveis: limite de proporcionalidade, limite de escoamento, limite de resistência à tração e a ruptura.

Fonte: os autores

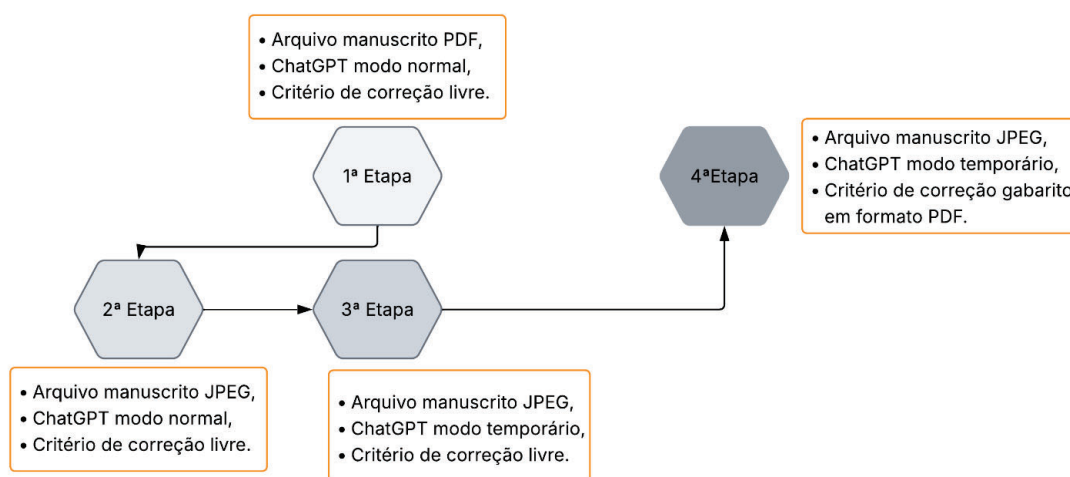
Figura 4 - Questão discursiva textual

7) (Valor 1,6 pontos) Um determinado engenheiro tem como objetivo produzir um eixo de transmissão utilizando aço AISI 1045. Porém, ele verifica que esse material não tem a resistência mecânica necessária para a aplicação desejada. Diante disso, o que você indicaria para o engenheiro para aumentar a resistência mecânica do material sem modificar a sua microestrutura? Justifique a sua resposta aplicando os conceitos estudados.

Fonte: os autores

O percurso metodológico foi estruturado em quatro etapas principais, visando testar diferentes estratégias de interação com a ferramenta de IA (ChatGPT - 4.0) e avaliar sua capacidade de analisar e corrigir avaliações manuscritas, conforme descrito na Figura 5.

Figura 5 - Fluxograma das etapas dos procedimentos da análise



Fonte: os autores

Na primeira etapa, foi inserido diretamente na plataforma um arquivo PDF contendo uma prova manuscrita. O comando utilizado solicitava uma correção livre, com atribuição de nota de 0 a 10, sem critérios previamente definidos.

Na segunda etapa, continuamos avaliando a viabilidade do método empregado para anexar os arquivos. Nesta etapa os arquivos foram convertidos para o formato JPEG, a fim de verificar se esse tipo de extensão melhoraria a extração de informações. Essas duas etapas iniciais consideramos a viabilidade do método, no sentido de verificar qual o melhor formato de arquivo para anexar no ChatGPT.

Na terceira etapa, repetiu-se o teste com os arquivos em formato JPEG, mas desta vez utilizando o modo de chat temporário da IA, o qual não mantém o histórico de conversas anteriores. Essa abordagem visou minimizar ruídos contextuais e melhorar a objetividade das respostas.

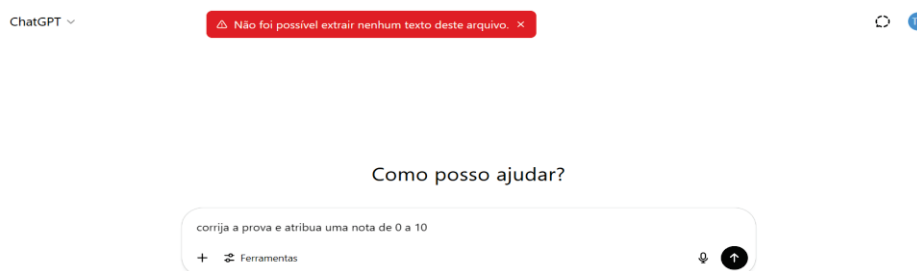
Por fim, na quarta etapa, buscou-se orientar a ferramenta de maneira mais precisa, por meio do fornecimento de um gabarito estruturado elaborado pelos próprios professores. Esse documento, originalmente em formato Word, foi convertido em PDF e inserido na plataforma juntamente com a prova em JPEG. O *prompt* utilizado solicitava expressamente a correção da avaliação com base no gabarito, respeitando os critérios e pontuações ali descritos. Essas duas etapas tiveram como objetivo analisar a precisão da IA na correção das avaliações.

## 4. Resultados obtidos

### 4.1. Viabilidade do Método

Inicialmente anexamos as provas manuscritas pelos alunos digitalizadas em formato PDF no site do ChatGPT. Com esse tipo de formato não se obteve sucesso, a ferramenta retornou a mensagem: “Não foi possível extrair nenhum texto do arquivo”, conforme mostrado na Figura 6. Isso evidenciou uma limitação relevante da IA na leitura de textos manuscritos presentes em arquivos PDF.

Figura 6 - Print da tela com a mensagem fornecida pela ferramenta



Fonte: Tela site ChatGPT

Diante disso, optou-se por digitalizar as provas e salvá-las em arquivo JPEG, etapa 2. Com as imagens digitalizadas, a ferramenta conseguiu realizar a leitura do conteúdo e gerar correções. Contudo, observou-se a ocorrência de respostas completamente desconexas das questões apresentadas, fenômeno conhecido como “alucinação” da IA. As inconsistências foram identificadas a partir da análise crítica dos próprios *feedbacks* produzidos pela ferramenta. Sendo assim, constatamos que o formato de arquivo JPEG era apropriado, mas precisávamos eliminar o problema da “alucinação”. Para tentarmos solucionar o problema fizemos a terceira etapa conforme descrito no fluxograma da Figura 5, empregando o modo temporário do ChatGPT. Essa ideia foi uma boa estratégia porque ele não mantém o histórico de conversas anteriores. Com essa abordagem ocorreu uma minimização dos ruídos contextuais e melhorou a objetividade das respostas.

Logo ressaltamos que a utilização de arquivos em formato JPEG, em substituição ao PDF, mostrou-se eficaz para viabilizar a leitura do conteúdo manuscrito, permitindo a continuidade das análises. Além disso, o uso do chat temporário reduziu ruídos contextuais que poderiam comprometer a imparcialidade das respostas, aumentando a confiabilidade das correções realizadas.

#### 4.2. Precisão da Correção

As provas foram elaboradas com questões diversificadas tanto na disciplina de Estrutura dos Materiais como nas de Propriedades dos Materiais, conforme mencionado no percurso metodológico. Na Tabela 1 apresentamos a relação dos tipos de questões contidas nas avaliações e a porcentagem de precisão dos acertos da IA para cada etapa do processo.

Tabela 1 - Avaliação da precisão de correção da IA para os formatos de questões avaliados

<b>Tipo de Questão</b>	<b>Quantidade de Questão</b>	<b>Etapa 1</b>	<b>Etapa 2</b>	<b>Etapa 3</b>	<b>Etapa 4</b>
Multipla escolha textual	5	0%	0%	20%	60%
Multipla escolha com imagens	2	0%	0%	50%	50%
Discursiva com interpretação de gráficos e/ou tabelas	1	0%	0%	0%	100%
Discursiva com solicitação de desenho técnico	2	0%	0%	50%	50%
Discursiva textual	2	0%	0%	70%	70%

Fonte: os autores

Observamos na Tabela 1 que, nas etapas 1 e 2, a IA não foi eficaz na correção das avaliações, apresentando 0% de precisão. Esse fato é resultante das conclusões anteriores. Como já indicado, a ferramenta teve dificuldades em extrair texto de arquivos PDF manuscritos, o que exigiu a conversão para imagens no formato JPEG. Outro ponto crítico foi a ocorrência de alucinações, as respostas geradas não tinham relação com as perguntas apresentadas, o que comprometeu a validade de algumas correções.

Na terceira etapa observamos um desempenho inconsistente na correção das questões de múltipla escolha que envolviam apenas textos, 20% de precisão acertando 1 de 5 questões analisadas. Apesar da persistência de inconsistências, a precisão melhorou com a inserção do gabarito, atingindo 60%. Acreditamos que essa precisão melhorou, mas ainda não é de 100% por conta da dificuldade do ChatGPT em ler arquivos no formato JPEG.

Em relação às questões de múltipla escolha que envolviam gráficos ou imagens a porcentagem de acertos foi de 50% (1 de 2) com ou sem a inserção do gabarito. Já para o caso da questão discursiva com interpretação de gráficos e/ou tabelas a IA inicialmente não acertou a correção, mas ao comparar com o gabarito ela consegue corrigir corretamente. De acordo com Santos (2023) o

ChatGPT apresentava limitações significativas. Sua taxa de acerto varia entre 13,6% e 55,5%, considerando problemas de interpretação (especialmente em gráficos e formas geométricas), raciocínios incoerentes, cálculos incorretos, e dificuldade com representações visuais abstratas. Além disso, o modelo comete erros primários, como ignorar dados relevantes ou chegar ao valor correto e escolher a alternativa errada. Em alguns casos, recusa-se a responder por “falta de informações”. Logo, o ChatGPT ainda enfrenta desafios consideráveis, especialmente na lógica, precisão de cálculos e interpretação de elementos visuais, indicando necessidade de aprimoramento (Santos, 2023; Celestino e Valente, 2024).

Contudo acreditamos que esse problema também é visto em nossas disciplinas, no que diz respeito à interpretação de elementos visuais e ao conhecimento específico relacionado ao conteúdo técnico explorado. Isso nos levou a concluir que a inteligência possui a necessidade de aprimoramento em relação aos conteúdos de Estrutura dos Materiais e Propriedade dos Materiais. Como o ChatGPT é uma inteligência artificial baseada em um modelo de linguagem natural que permite simular conversas humanas, há a necessidade de treiná-lo com um grande conjunto de dados, permitindo que ele aprenda e evolua a cada interação. Esse fato foi visto no uso do ChatGPT na resolução da prova da OAB que demonstrou como o treinamento constante pode aprimorar sua capacidade de análise e precisão (Inteligência artificial é aprovada em 1ª fase da OAB, 2024).

Para às questões discursivas com solicitação de desenhos técnicos a IA obteve 50% de acertos (1 de 2), quando usado o critério da terceira etapa. Esse valor permaneceu mesmo utilizando o gabarito como referência. Mais uma vez tivemos o indício da dificuldade da inteligência em interpretar imagens.

Por fim, em relação às questões somente discursivas o ChatGPT conseguiu corrigi-las com um grau de precisão maior, não houve a necessidade de leitura de gráficos ou ilustrações, vemos na Tabela 1 uma porcentagem de 70%. Esse valor está relacionado com a superficialidade das respostas dadas pela IA. Isso foi identificado a partir dos próprios *feedbacks* gerados pela ferramenta. Portanto, mesmo que o aluno discorra superficialmente a inteligência considera como correta sua resposta e atribui nota máxima à questão, muito diferente daquela atribuída pelo docente.

Sendo assim, esse conjunto de estratégias permitiu avaliar, em diferentes níveis, a incapacidade do ChatGPT em atuar na correção de provas manuscritas e com conteúdo técnicos complexos. Mesmo com a inserção de um gabarito estruturado, as limitações persistiram, o que reforça a necessidade de mediação humana qualificada na interpretação e aplicação dos critérios avaliativos. Acreditamos que isso pode ser melhorado ao passo que a inteligência seja treinada de acordo com os conteúdos

técnicos, entretanto, é crucial considerar a dificuldade em relação a extração de imagens. Outra observação é uma tendência em atribuir notas superiores às definidas pelos docentes, indicando falta de rigor na análise de respostas discursivas mais complexas.

Assim, conforme destaca o Professor Preguiça (2024), a inteligência artificial pode ser uma ferramenta útil e complementar ao trabalho docente, mas ainda está longe de assumir de forma autônoma a responsabilidade de corrigir avaliações em contextos educacionais exigentes como o do ensino superior em Engenharia.

## **5. Lições aprendidas e direções para estudos ou aplicações futuras**

A realização desta experiência trouxe importantes lições sobre o potencial e os limites do uso do ChatGPT na educação, especialmente em contextos que envolvem avaliação de conteúdos técnicos complexos, como é o caso das disciplinas do curso de Engenharia.

A primeira lição refere-se à necessidade de adaptação dos formatos de arquivos, uma vez que o ChatGPT demonstrou maior desempenho com imagens do que com documentos PDF manuscritos. Além disso, ficou evidente que a clareza dos comandos (*prompts*) influencia diretamente a qualidade das respostas, exigindo do usuário um domínio mínimo da lógica de interação com a ferramenta.

Outra lição fundamental diz respeito à importância da supervisão humana no processo avaliativo. Embora a IA possa ser uma ferramenta de apoio para tarefas repetitivas ou correções preliminares, ela não é capaz de interpretar nuances de raciocínio, criatividade ou domínio técnico da mesma forma que um professor experiente. A ferramenta tende a seguir padrões genéricos de correção e demonstra dificuldades em compreender o contexto específico de cada questão, principalmente quando envolve linguagem simbólica, representação gráfica ou elaboração de desenhos técnicos.

Baseado nos resultados, pode-se observar que o uso da inteligência artificial como aliada no processo de avaliação é promissor, desde que compreendido como complementar à atuação docente. A experiência demonstrou que, com ajustes metodológicos e mediação pedagógica, é possível integrar essa tecnologia de forma produtiva ao cotidiano acadêmico, otimizando etapas operacionais da correção sem comprometer a qualidade e a profundidade da avaliação. Essa integração, no entanto, exige formação contínua dos docentes, reflexão crítica sobre as limitações dos algoritmos e responsabilidade ética na aplicação de novas tecnologias educacionais.

## Agradecimentos

Agradecemos ao apoio institucional do Programa de Inovação e Assessoria Curricular (PROAIC) e do Departamento de Engenharia Metalúrgica e de Materiais (VMT) da Escola de Engenharia Industrial Metalúrgica de Volta Redonda (EEIMVR) - Universidade Federal Fluminense, cuja colaboração foi essencial para o desenvolvimento e a realização deste trabalho.

## Referências

CARVALHO, Joana D’Arc Valgueiro Barros; LINS, Micherllayne Alves Ferreira; GUSMÃO, Diógenes. **Adoecimento psíquico de professores: análise e reflexões a partir da literatura**. Revista FT, Linguística, Letras e Artes, v. 29, 2025. DOI: <https://doi.org/10.69849/revistaft/ar10202505031741>

CELESTINO, Marcelo Salvador; VALENTE, Vania Cristina Pires Nogueira. **O uso da ferramenta ChatGPT no suporte à educação e à produção acadêmica**. Educação Temática Digital, Campinas, v. 26, p. 1–20, 2024. DOI: <https://doi.org/10.20396/etd.v26i00.8673464>

INTELIGÊNCIA artificial é aprovada em 1ª fase da OAB. Estratégia OAB, 17 jun. 2024. Disponível em <<https://oab.estrategia.com/portal/inteligencia-artificial-e-aprovada-em-1a-fase-da-oab/>> Acesso em: 28 maio. 2025.

LIMA, Fábio Gotz de; PALOSK, Luís Henrique. **Saúde mental e estilos de ensino de professores universitários de instituições públicas e privadas**. Boletim de Conjuntura – BOCA, Boa Vista, v. 17, n. 51, 2024. DOI: <https://doi.org/10.5281/zenodo.10895714>

MOURA, J. S. et al. **A precarização do trabalho docente e o adoecimento mental no contexto neoliberal**. Revista Profissão Docente, v. 19, n. 40, 2019. DOI: <https://doi.org/10.31496/rpd.v19i40.1242>

OKAGBUE, E. F. et al. **A comprehensive overview of artificial intelligence and machine learning in education pedagogy: 21 years (2000–2021) of research indexed in the Scopus database**. Social Sciences & Humanities Open, v. 8, 100655, 2023.

OPENAI. How ChatGPT and our foundation models are developed. 2023. Disponível em: <<https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed>> Acesso em: 10 maio. 2025.

PROFESSOR PREGUIÇA. **Inteligência artificial na correção de provas: como ela funciona e onde ainda erra**. Blog Professor Preguiça, 24 abr. 2024. Disponível em: <<https://blog.professorpreguica.com.br/posts/inteligencia-artificial-na-correcao-de-provas>> Acesso em: 28 maio. 2025.

SANTOS, Juliano Nunes dos. **Avaliação do desempenho do ChatGPT na área de Matemática dos testes do ENEM. 2023**. Monografia (Graduação em Ciência da Computação) – Centro de Informática, Universidade Federal da Paraíba, João Pessoa, 2023. Disponível em: <[https://repositorio.ufpb.br/jspui/handle/123456789/31608?locale=pt\\_BR](https://repositorio.ufpb.br/jspui/handle/123456789/31608?locale=pt_BR)> Acesso em: 10 maio 2025.

SANTOS, N. C.; CORTEZ, E. A. **Educação permanente como estratégia de promoção da saúde mental para docentes universitários**. Research, Society and Development, v. 10, n. 8, 2021. DOI: <https://10.33448/rsd-v10i8.17235>

WAGNER, T. F. et al. **Ansiedade social e comorbidades em professores do ensino superior**. Revista Psicologia e Saúde, 2021.