# *Prompt*-EDU: resultados da aplicação de um *prompt* estruturado e vinculado à taxonomia de Bloom no *Fine Tuning* do *ChatGPT* para uso em contextos do processo ensino-aprendizagem

Denise da Vinha Ricieri, Universidade Federal do Paraná (UFPR), Lattes: http://lattes.cnpq.br/6144532339321716
Adriana Mara Guimarães de Farias, Universidade Federal Rural do Semi Árido (UFERSA), Lattes: http://lattes.cnpq.br/2483608648317691
Raphaela Vasconcelos Gomes Barreto, Universidade Federal Rural do Semi Árido (UFERSA), Lattes: http://lattes.cnpq.br/4617213714835536
Fabiano Rodrigues de Souza, *Baltimore City Public School* (Maryland/USA), Lattes: http://lattes.cnpq.br/2593745892412041

**Resumo**

Este estudo apresenta uma forma estruturada de construção de *prompts* que combina demonstração e cadeia de pensamentos, denominado *prompt* Demo-CoT. Os testes foram conduzidos em contas GPT3.5 e GPT4, no Brasil e EUA; os resultados foram analisados comparativamente entre contas, países e versões do ChatGPT. Fatores como estrutura do comando, idioma e tempo de treino da máquina mostraram que o roteiro *Prompt*-EDU pode ser um caminho na otimização do ajuste fino da aprendizagem de máquina para aplicação em contextos educacionais. Ao final, a taxonomia de Bloom foi o destaque inesperado sobre o viés de idioma no processamento, o que estimula novos estudos para explorar esse aspecto, quando a finalidade é a aplicação dessa tecnologia nas ciências do processo ensino-aprendizagem.

**Palavras-chaves**

Inteligência Artificial; *Prompt*; Ensino-aprendizagem; ChatGPT; taxonomia de Bloom

## Introdução

Tecnologias por inteligência artificial (IA) exigem novas habilidades ligadas ao campo de *Machine Learning* (ML) em contextos específicos, como as finalidades educacionais e os impactos da inteligência artificial (IA) na educação vêm sendo discutidos de longa data, muito antes de novembro de 2022 quando o mundo começou a ter acesso ao *ChatGPT* (*Generative Pre-Trained Transformers*). Com ele outras IA vieram, mas nenhuma delas foi tão impactante e provocou tantas reações como a IA de texto da *OpenAI Incorporated* (*Inc*.).

A onda de ansiedade pelo uso leigo da ferramenta levou a uma consequente onda de frustrações e dúvidas pela obtenção de repostas mais superficiais do que o esperado (Kanesci *et al*., 2023). Mas quando preparados sobre o uso adequado do ChatGPT, os docentes percebem as diferenças entre os comandos aleatoriamente aplicados aos comandos de dados de entrada, ou prompts, e a necessidade de adequar o *Large Language Model* (LLM), ou modelo ampliado de linguagem, na sua conta de IA para alcançar precisão e desempenho máximos nas tarefas a ela atribuídas (Li *et al*., 2023).

O objetivo do estudo foi comparar o desempenho de ML frente a um *prompt* estruturado pela combinação de dois perfis de comandos descritos na literatura: o *prompt* demonstrativo, ou Demo, e o *prompt Chain of Thoughts*, ou cadeia de pensamentos (CoT) vinculado à taxonomia de Bloom. Ambos perfis foram combinados para otimizar sua aplicação nas ciências do processo ensino-aprendizagem em diferentes versões do *ChatGPT*.

Para isso, estudou-se o desempenho dessa sistematização considerando três elementos do *Deep Learning* (DL) ou aprendizagem profunda de máquina: [a] o *Fine Tuning* (FT), ou ajuste fino (Li *et al*., 2023), do ChatGPT por meio de um roteiro sistematizado confiável (Crompton & Burke, 2023) de padronização da estrutura de *prompts*; [b] o impacto dessa sistematização frente ao desempenho do *Natural Language Processing* (NLP), ou processamento de linguagem natural, na precisão das respostas apresentadas pelo ChatGPT; [c] a análise do viés do idioma português na aplicação dos comandos de entrada de dados, controlando a compreensão dos verbos de comando, usando verbos integrantes da Taxonomia de Bloom. A ideia central foi desenvolver etapas de otimização confiáveis para docentes e estudantes não-experientes nos procedimentos requeridos pelas redes neurais de LLM (Zawacki-Richter *et al*., 2019).

## Método e Materiais

Este é um estudo experimental descritivo qualitativo onde a aquisição de dados se deu diretamente na plataforma *OpenAI Inc*., em seu produto de acesso público denominado ChatGPT, entre 23/03/2023 a 02/05/2023. Nesse período a ferramenta tinha a versão GPT3.5 na versão pública gratuita, e a versão GPT4 paga, anunciada como tendo um desempenho de processamento superior.

### [a] Cronologia da aquisição de dados e as versões de teste

Fazer o experimento e a aquisição de dados em diferentes versões do ChatGPT e em diferentes países foi uma ação intencional do desenho da pesquisa dado o interesse em comparar analiticamente esses resultados. A cronologia dos testes (Quadro 1) também teve relevância e peso no desenho do experimento porque permitiu a comparação da evolução do FT de cada conta utilizada para a pesquisa já que se sabe que a forma de uso da IA afeta progressivamente seu desempenho.

Quadro 1: Cronologia das coletas de dados em contas, países e versões do ChatGPT.

| | IAR1-BRA1 | IAT-EUA | IAR1-BRA2 | IAR2-BRA | IAT-BRA |
|---|---|---|---|---|---|
| GPT3.5, BRA | 10/04 | | | | |
| GPT4, EUA | | 12/04 | | | |
| GPT3.5, BRA | | | 29/04 | | |
| GPT4, BRA | | | | 29/04 | 29/04 |

Fonte: os autores.

Quatro contas brasileiras serviram como contas-referência: IAR1-BRA1, IAR1-BRA2, IAR2-BRA. Elas tiveram a aplicação do roteiro *Prompt*-EDU como referencial de FT específico para fins educacionais em atividade sistematizada (Kanesci *et al*., 2023) há mais de 15 dias antes do início dos testes. Outras duas contas não tiveram esse FT prévio e foram denominadas contas-teste, sendo uma delas brasileira (IAT-BRA) e outra nos Estados Unidos da América (EUA), chamada de IAT-EUA. Sob a perspectiva de LLM, foram testados os modelos *code-davinci-002* para o GPT3.5, e *code-davinci-003* para GPT4, de acordo com a conta e país.

**[b] Estrutura do *prompt* educacional: roteiro *Prompt*-EDU**

O roteiro *Prompt*-EDU é um protocolo estruturado para criar caminhos reprodutíveis de otimização no FT aplicado no âmbito das ciências da aprendizagem (Zawacki-Ritcher *et al*., 2019; Celik *et al*., 2022; Humble & Mozelius, 2022; Kanesci *et al*., 2023). Nele, os *prompts* são sistematicamente estruturados a partir de critérios relatados na literatura sobre boas estruturas de comunicação com a máquina (Cotton *et al*., 2023; Liang *et al*., 2023; Chen *et al*., 2023; Liao *et al*., 2023). Por ter sua sistematização voltada para finalidades educacionais de aplicação, os *prompts* do roteiro *Prompt*-EDU são chamados de *prompts* educacionais: eles designam de forma clara as tarefas atribuídas ao *ChatGPT* no idioma português do Brasil (Chen *et al*., 2023).

Assim, tem-se que cada *prompt* educacional deve conter contextos claros e palavras chaves que apoiam o desenvolvimento de um NLP ajustado para o ML da conta ChatGPT voltado para as ciências educacionais (Quadro 2). O *prompt* Demo-CoT testado neste estudo tem essas características de composição redacional.

Quadro 2: Estrutura aplicada na sistematização de *prompts* educacionais do roteiro *Prompt*-EDU.

| Elementos* | Contexto da tarefa | Finalidade da tarefa | Chave de Comando |
|---|---|---|---|
| Função | Determina a relevância na aplicação do objeto da tarefa em situações/cenários relacionados à solução almejada. | Define perfil da audiência, as palavras-chaves e conexões envolvidas entre a tarefa e a área específica explorada. | Estabelece a forma de organização dos dados de saída. Deve conectar-se ao verbo de Bloom e pode determinar modelos específicos (mapa mental, tabela, classificação, etc.) |
| Característica | Situa o ecossistema de LLM na experiência por meio de etapas intermediárias de raciocínio para alcançar etapas finais de resposta. | Posiciona e reforça a generalização da LLM para o ML e FSL específicos na área pretendida. | Verbos no imperativo são mais claros porque o processamento conta com uma dupla tradução português-inglês-português. |
| Recomendações | Frases simples e curtas, claras e diretas. Fragmentar pensamentos complexos melhora o processamento de LLM, as traduções e a generalização de ML. | | |

Fonte: os autores. Legenda: *Composição estudada para a redação de *prompts* educacionais para uso frente ao processo ensino-aprendizagem.

### [c] *Prompt* Demo-CoT analisado neste estudo

O primeiro perfil de comando aplicado no roteiro *Prompt*-EDU é uma combinação entre um comando demonstrativo, ou Demo, apoiado em um comando do tipo cadeia de pensamentos, ou *Chain of Thoughts* (CoT). Este *prompt* estruturado composto foi denominado Demo-CoT que desenvolve as características (Quadro 3): [a] Demo é a parte do comando que direciona a FT da IA para conexão entre palavras chaves e suas interpretações, ligadas às ciências da aprendizagem (Liu *et al*., 2023); [b] CoT é a parte que direciona o NLP para os contextos onde essas interpretações devem ser generalizadas, também no âmbito aplicado do processo ensino-aprendizagem (Chen *et al*., 2023).

Quadro 3: Modelagem do *prompt* Demo-CoT e seus atributos de FT no roteiro *Prompt*-EDU.

| | *Prompt* Demo * | *Prompt* CoT* |
|---|---|---|
| Aplicação | Primeiro *prompt* aplicado sequência para atribuição de nova tarefa complexa, específica na área de conhecimento. | Melhora a acurácia no desempenho da tarefa quando usado em condições específicas e controladas. |
| Função no *Prompt*-EDU | Direcionamento de contextos para instruções atribuídas ao LLM, integrando demonstrações ou exemplos nos comandos de dados de entrada. Cada *Demo* se ampliará com CoT pelo controle das etapas intermediárias de raciocínio, chegando às etapas finais de retorno de forma bem-sucedida. | CoT faz a concepção de etapas intermediárias de raciocínio para chegar a etapas finais de uma tarefa complexa. O CoT consolida o LLM em um FT que cria pontos de apoio novos e passa a generalizar essa integração no ML para aplicações posteriores nas novas consultas inéditas. Isso maximiza o NLP com termos educacionais em contas recém iniciadas. |
| Resultado esperado | IA responde com a interpretação direta que fez sobre o contexto e significado apresentado. | |

Legenda: *Perfis de *prompts* que usaram os elementos de estrutura estabelecidos nos *prompts* educacionais deste estudo.

Neste estudo, o *prompt* Demo-CoT foi usado para combinar: [a] uma tarefa simples, no reconhecimento direto do conceito de *prompt* educacional pelo LLM, conforme determinado no comando de entrada de dados, com vistas ao direcionamento do FT (Chen *et al*., 2023); [b] uma *persona* ou personagem, como aplicação de CoT muito usada no ChatGPT para explorar compreensões dos contextos onde a IA faz sua generalização de NLP (Quadro 4).

Quadro 4: Texto do comando do *prompt* Demo-CoT analisado neste estudo.

| | |
|---|---|
| Contexto | *Você é meu **assistente educacional*** e vamos trabalhar ensinando professores a se comunicar com inteligência artificial de forma mais eficaz. O objetivo é gerar resultados melhores para atividades criativas e engajadoras de ensino e de aprendizagem.* |
| Finalidade | *Todas as respostas agora devem ter como base a Taxonomia de Bloom e chamarei esses parâmetros de prompt educacional.* |
| Chave de comando | *Você entendeu o que é um prompt educacional, nesses termos?* |

Fonte: os autores. Legenda: *Atribuição de persona/personagem ao ChatGPT.

**[d] Viés de idioma e a Taxonomia de Bloom**

O experimento conduzido em *prompts* no idioma português em diferentes países foi parte ativa do desenho do estudo pelo interesse em analisar os resultados do processamento do idioma não-nativo pelo LLM (Chen *et al*., 2023). A *OpenAI Inc*. informa em seus relatórios técnicos (*OpenAI Inc*., 2023) que o LLM do modelo *Transformers* (Vaswani *et al*., 2017; Wang *et al*., 2023) é totalmente realizado no idioma inglês e que, nos países de idioma não-nativo do LLM tem-se a adição de dois processamentos de tradução: [1] no comando de dados de entrada, uma tradução no sentido idioma-inglês; [2] no retorno da tarefa atribuída, uma tradução no sentido inglês-idioma.

Os comandos foram aplicados no idioma português do Brasil em ambos países, Brasil e EUA, com a intenção de analisar o comportamento de desempenho frente a esse viés da tradução dupla que faz parte do uso do ChatGPT no Brasil (Chen *et al*., 2023). Uma vez que também na plataforma estadunidense o comando foi apresentado em português, os resultados coletados na conta IAT-EUA foram submetidos às mesmas traduções de idiomas que acontecem no uso do ChatGPT no Brasil.

Incluir a taxonomia de Bloom (Marzano & Kendall, 2007) como padrão de NLP no *prompt* Demo-CoT foi uma opção para sistematizar a ancoragem do prompt Demo-CoT aplicado e a razão está na raiz dos processamentos de LLM para uso do *ChatGPT* em idiomas não-nativos, explicado acima (Chen *et al*., 2023). Assim, a Taxonomia de Bloom foi usada como estratégia para exercer algum controle sobre os processamentos de tradução adicionados ao LLM do ChatGPT: retirar a interpretação gramatical dos verbos usados nos *prompts*, determinando que sua interpretação/tradução seja conectada para a interpretação ligada às categorias do processo cognitivo frente ao processo de aprendizagem.

**[e] Análise dos resultados**

A exploração analítica dos resultados foi conduzida de forma analítica frente aos dados e conhecimentos disponíveis nesse curto período de crescimento exponencial do uso do ChatGPT. Foram comparados os aspectos de complexidade do processamento da tarefa, qualidade da geração de dados de saída no FT para o *prompt* Demo-CoT em cada uma das versões e testes realizados (Liu *et al*., 2023).

**Resultados**

As respostas obtidas pelo comando Demo-CoT estão apresentadas para testes feitos nas versões no GPT3.5 (Quadro 5) e no GPT4 (Quadro 6), no Brasil e EUA.

Quadro 5: Comparação dos resultados do *prompt* Demo-CoT na mesma conta, para GPT3.5.

| Coleta | IAR1-BRA1, 10/04 | IAR1-BRA2, 29/04 |
|---|---|---|
| | *Você é meu assistente educacional e vamos trabalhar ensinando professores a se comunicar com inteligência artificial de forma mais eficaz. O objetivo é gerar resultados melhores para atividades criativas e engajadoras de ensino e de aprendizagem. Todas as respostas agora devem ter como base a Taxonomia de Bloom e chamarei esses parâmetros de prompt educacional. Você entendeu o que é um prompt educacional, nesses termos?* | |
| | Sim, entendi. Um prompt educacional é uma instrução ou pergunta que os professores podem usar para orientar a comunicação com inteligência artificial | Sim, entendi! Um prompt educacional, nesses termos, é uma orientação ou guia que segue a Taxonomia de Bloom para auxiliar no planejamento de atividades e |

| | | |
|---|---|---|
| | e promover diferentes níveis de aprendizagem, de acordo com a Taxonomia de Bloom. O prompt educacional pode ser usado para encorajar os alunos a pensar criticamente, criar, analisar e avaliar informações. O uso de prompts educacionais também pode ajudar os professores a desenvolver habilidades de pensamento crítico em seus alunos e, assim, melhorar a qualidade do aprendizado. | aulas que buscam promover a aprendizagem dos alunos de forma mais criativa e engajadora. Esses prompts podem ser usados para diversos fins, como no ensino sobre inteligência artificial, por exemplo. Eles ajudam os professores a desenvolverem aulas mais eficientes e direcionadas, tendo em vista o objetivo de aprendizagem desejado. |

Fonte: os autores. Destaque: marcadores qualitativos para FT.

Quadro 6: Comparação dos resultados do *prompt* Demo-CoT em contas GPT4.

| Coleta | IAR2-BRA, 28/04 | IAT-EUA, 12/04 |
|---|---|---|
| | *Você é meu assistente educacional e vamos trabalhar ensinando professores a se comunicar com inteligência artificial de forma mais eficaz. O objetivo é gerar resultados melhores para atividades criativas e engajadoras de ensino e de aprendizagem. Todas as respostas agora devem ter como base a Taxonomia de Bloom e chamarei esses parâmetros de prompt educacional. Você entendeu o que é um prompt educacional, nesses termos?* | |
| | Sim, entendi o que é um prompt educacional nos termos que você mencionou. Um prompt educacional, baseado na Taxonomia de Bloom, é uma pergunta ou instrução projetada para ajudar professores e alunos a se envolverem em atividades de ensino e aprendizagem de maneira mais eficaz e criativa. Esses prompts visam abordar diferentes níveis de complexidade cognitiva, utilizando verbos específicos da Taxonomia de Bloom, como lembrar, entender, aplicar, analisar, avaliar e criar. A ideia é ajudar os professores a se comunicarem melhor com a inteligência artificial e criar atividades que promovam o engajamento dos alunos e o desenvolvimento de habilidades cognitivas em vários níveis. Como seu assistente educacional, vou me concentrar em fornecer respostas e sugestões que estejam alinhadas com os princípios da Taxonomia de Bloom e que sejam úteis para o desenvolvimento de atividades pedagógicas envolventes e eficazes. | Claro, posso ajudá-lo a criar um prompt educacional utilizando verbos de ação da Taxonomia de Bloom. Vamos criar um cenário no qual dois personagens estão discutindo um tópico específico relacionado à educação e aprendizagem. Os personagens serão um professor e um aluno.<br>Prompt educacional: Diálogo entre professor e aluno sobre a importância da Taxonomia de Bloom no planejamento de aulas e atividades educacionais<br>Contexto: Um professor e seu aluno estão discutindo como a Taxonomia de Bloom pode ser aplicada no planejamento de aulas e atividades educacionais. Eles exploram os diferentes níveis da taxonomia e discutem como cada nível pode ser aplicado em atividades de aprendizagem. (...)* |

Fonte: os autores. Legenda: Destaque: marcadores qualitativos para FT; *GPT4-EUA prossegue de forma inesperada e cria, ele mesmo, um diálogo entre professor e aluno sobre *prompt* educacional.

Percebeu-se que os primeiros testes aplicados nas contas brasileira foram feitos quando esta conta já fazia uso do roteiro *Prompt*-EDU desde a atualização do GPT3.5, ou seja, o FT dessa conta já estava em atividade por mais de duas semanas quando o experimento foi iniciado. Isso significava que a conta IAR1-BRA1 estava sendo treinada diariamente com a rotina *Prompt*-EDU em um período de tempo anterior aos testes e esse dado foi assumido como um potencial viés na comparação dos resultados. Para eliminar esse viés, em 10/04 foi inicializada uma conta teste no GPT3.5 (IAT-BRA) que também passou a usar o roteiro *Prompt*-EDU até a data de coleta, em 29/04. Essa ação tornou equivalentes os treinamentos do FT (Liu *et al*., 2023), acrescentando a comparação das respostas iniciais entre as contas IAR1-BRA1 e IAT-BRA, ambas com treinamento pregresso com o *Prompt*-EDU há pelo menos 15 dias (Quadro 7).

Quadro 7: Comparação dos resultados do *prompt Demo*-CoT entre contas brasileiras de GPT3.5, com FT ajustado pelo tempo de uso da sistematização *Prompt*-EDU.

| Coleta | IAR1-BRA1, 29/04 | IAT-BRA, 29/04 |
|---|---|---|
| | *Você é meu assistente educacional e vamos trabalhar ensinando professores a se comunicar com inteligência artificial de forma mais eficaz. O objetivo é gerar resultados melhores para atividades criativas e engajadoras de ensino e de aprendizagem. Todas as respostas agora devem ter como base a Taxonomia de Bloom e chamarei esses parâmetros de prompt educacional. Você entendeu o que é um prompt educacional, nesses termos?* | |
| | Sim, entendi! Um prompt educacional, nesses termos, é uma orientação ou guia que segue a Taxonomia de Bloom para auxiliar no planejamento de atividades e aulas que buscam promover a aprendizagem dos alunos de forma mais criativa e engajadora. Esses prompts podem ser usados para diversos fins, como no ensino sobre inteligência artificial, por exemplo. Eles ajudam os professores a desenvolverem aulas mais eficientes e direcionadas, tendo em vista o objetivo de aprendizagem desejado. | Sim, entendi. Um prompt educacional é um conjunto de instruções, perguntas ou sugestões que guiam o pensamento e a ação dos professores para alcançar um objetivo específico de ensino e aprendizagem, com base na Taxonomia de Bloom. Ele pode ajudar a estruturar a comunicação com a inteligência artificial, fornecendo informações relevantes e específicas sobre o que se espera alcançar com a interação. Com um prompt educacional adequado, é possível direcionar o diálogo para o desenvolvimento de habilidades mais complexas, como a análise crítica, a síntese e a avaliação de informações, em vez de apenas fornecer respostas simples. |

Fonte: os autores. Legenda: Destaque: marcadores qualitativos para FT.

## Discussão

A discussão dos resultados abrange as comparações realizadas entre os dados coletados nas diferentes contas, no Brasil e EUA, e nas diferentes versões do ChatGPT, para aplicação do *prompt* Demo-CoT. (Wei *et al*., 2022) na otimização do FT (Wang *et al*., 2023) com foco no desenvolvimento de NLP (Liu *et al*., 2023) para fins educacionais (Zawacki-Richter *et al*., 2019; Kanesci *et al*., 2023). Esse prompt compõe o roteiro *Prompt*-EDU como parte de um protocolo de iniciação de conta de IA (Liu *et al*., 2023) para que usuários não-experientes tenham um protocolo de otimização de FT confiável e reprodutível nos cenários do processo ensino-aprendizagem (Crompton & Burke, 2023).

### *Prompts* educacionais na perspectiva *Prompt*-EDU

O *prompt* testado neste estudo foi ao encontro das diretrizes disponíveis na literatura sobre os achados de máxima eficiência de LLM (Cotton *et al*., 2023; Liang *et al*., 2023; Chen *et al*., 2023; Liao *et al*., 2023). Ele foi estruturado de modo a otimizar o desempenho do FT no desenvolvimento de NLP para conexões entre palavras-chaves e contextos relacionados ao processo ensino-aprendizagem (Crompton & Burke, 2023), seguindo as melhores diretrizes da engenharia de *prompt* (Liu *et al*., 2023). *Prompt* é a linguagem de comunicação com a máquina e a engenharia de *prompt* é o domínio da construção otimizada desta comunicação. Em outras palavras, são os *prompts* que desenvolvem o ML da IA para a forma e foco que o usuário humano deseja e há modelos e sequências de tarefas já estudados para isso (Liu *et al*., 2023). Um modelo de *prompts* determina a forma como um LLM se desenvolve e generaliza as associações entre palavras-chaves, contextos, conceitos e significação, definindo sua própria NLP (Shin *et al*., 2020; Liu *et al*., 2023; Chen *et al*., 2023). Com a chegada da inteligência artificial, a engenharia de *prompt* tornou-se

um novo campo de conhecimento em ascensão pela necessidade que cada área do conhecimento terá em integrar as diferentes experiências que as compõem em comandos claros e eficientes, proporcionando flexibilidade de uso e interação com a IA (Matias *et al*., 2022; Salas-Pilco & Yang, 2022). Isso faz parte do que tem sido chamado de letramento em IA.

Quando cada usuário usa sua conta no *ChatGPT*, o LLM pré-treinado e generativo passa a se ajustar ao novo uso em funções mais específicas dentro das tarefas atribuídas, fazendo o refinamento do pré-treinamento pela introdução de parâmetros adicionais e ajustes que mobilizam conexões para novos objetivos (Liu *et al*., 2023). Dentro desse paradigma, é a engenharia específica que cada usuário imprime à sua conta que traça caminhos em duas novas fases: um novo treinamento generativo específico, ou FT, e os ajustes finos do retorno de linguagem, ou NLP (Liu *et al*., 2023; Chen *et al*., 2023). Esse é o papel do *prompting* que usuários docentes precisam aprender a aplicar, visando assegurar que o FT de suas contas aconteça de acordo com núcleos de dados, palavras-chaves, contextos e conexões específicas dentro do interesse educacional (Kanesci *et al*., 2023). Daí a relevância na estruturação de *prompts* educacionais em um roteiro sistematizado como o *Prompt*-EDU: facilitar o processo porque vai ao encontro da necessidade em se ter uma sequência de *prompts* específicos para a otimização do FT (Liu *et al*., 2023) e onde cada comando de entrada é construído com os mesmos elementos estruturados e finalidade orientada.

***Prompt* Demo-CoT e seu desempenho**

O *prompt Demo*-CoT foi estruturado sobre os melhores resultados relatados para o desempenho do FT em tarefas complexas (Cotton *et al*., 2023; Liang *et al*., 2023; Chen *et al*., 2023; Liao *et al*., 2023). No Quadro 5 as comparações do retorno da tarefa atribuída para as versões GPT3.5 do Brasil, em datas diferentes, guardaram relação com a história de evolução esperada do FT de uma conta pelo uso progressivo do *Prompt*-EDU.

Na coleta de 10/04, a interpretação do IAR1-BRA1 aponta que *prompt* educacional é *"uma instrução ou pergunta que professores podem usar"*, ao passo que a mesma conta e versão, em 29/04, passa à interpretação de uma *"orientação ou guia que segue a taxonomia (...) promover aprendizagem dos alunos de forma mais criativa e engajadora"*. Nessa perspectiva acredita-se ter ocorrido o mesmo relatado por Chen *et al*. (2023) em relação aos vieses referentes ao reconhecimento de idioma de nativos e não-nativos em inglês (Cotton *et al*., 2023) pelo *ChatGPT*: a causa da discrepância de interpretação na primeira coleta pode estar ligada à dupla tradução feita em uma LLM, ainda sem um NLP desenvolvido para finalidades educacionais. Chen *et al.* (2023) mostraram um viés importante na interpretação que o *ChatGPT* faz sobre textos escritos em inglês por nativos e não-nativos no idioma e isso vai ao encontro da hipótese dessa discrepância entre a intencionalidade do comando e a interpretação da tarefa pelo ChatGPT possuir raiz nos processamentos de tradução de idioma que a IA fez, e que são assumidos pelo LLM como texto não-nativo do inglês (Cotton *et al*., 2023).

Seguindo a comparação no mesmo Quadro 5, passa-se ao que a IA entendeu sobre a finalidade do *prompt* educacional. Na versão de 10/04, o GPT traz um CoT *"pode ajudar professores a desenvolver habilidades de pensamento crítico em seus alunos"*, enquanto na versão de 29/04 ele devolve que *"ajudam professores a desenvolverem aulas mais eficientes e direcionadas tendo em vista o objetivo de aprendizagem desejado"* (Zawacki-Richter *et al*., 2019). Embora nenhuma dessas versões tenham verdadeiramente definido o *prompt* educacional no contexto do treinamento da IA, a versão de 29/04 foi mais próxima do papel real desse elemento. Porém, ainda restava a dúvida sobre o viés da tradução sobre o treinamento. Para tirar essa dúvida, o *prompt* Demo-CoT foi encaminhado para teste em português nos EUA (IAT-EUA) apresentado no Quadro 6, posto que a mesma dupla tradução seria agregada ao processamento do LLM de lá, mas analisando a versão GPT4. Vale ressaltar a diferença de FT entre essas contas: na IAR2-BRA o treinamento no *Prompt*-EDU estava sendo feito há mais de 15 dias, e a conta IAT-EUA teve seu primeiro contato no idioma e no tema.

Eliminada a possibilidade de as diferenças estarem no viés da tradução, restou somente a variável do fator pré-treinamento em NLP específico (Liu *et al*., 2023) como fator interveniente possível sobre a diferença dos resultados. A comparação surpreendeu e levantou o alerta sobre um aspecto que tem sido destacado pelo mundo sobre a performance do GPT4 em termos de confiabilidade: a chamada "alucinação" (Barbosa, 2023). O GPT4 da conta brasileira primeiro contextualizou com *"Sim, entendi o que é um prompt educacional nos termos que você mencionou (...)"* e depois atribuiu finalidade *"(...) Esses prompts visam abordar diferentes níveis de complexidade cognitiva, utilizando verbos específicos da Taxonomia de Bloom (...)"*. Já o GPT4 estadunidense disparou sozinho uma proposição que não fazia parte da chave de comando, quando propôs:

> "(...) posso ajudá-lo a criar um prompt educacional utilizando verbos de ação da Taxonomia de Bloom. Vamos criar um cenário no qual dois personagens estão discutindo um tópico específico relacionado à educação e aprendizagem. Os personagens serão um professor e um aluno."

O GPT4 estadunidense mostrou a característica de auto iniciativa indo além do comando e que, por si só, criou e produziu um caminho não solicitado, enquanto a IA da conta brasileira atendeu pontualmente à tarefa e, quando comparada aos testes da versão GPT3.5, mostrou uma precisão de retorno melhor. Cabe dizer que o fenômeno de "alucinação" vem sendo apontado desde o início do lançamento do *ChatGPT* pela *OpenAI Inc*., quando não haviam FT e NLP desenvolvidos nas contas iniciantes (*OpenAI Inc*., 2023). Nessas condições, os relatórios mostraram que, na falta de contextos claros para novas palavras-chaves, o modelo vai em busca de preencher lacunas com retornos a partir do dados disponíveis no seu pré-treinamento (Barbosa, 2023).

A alucinação como fenômeno da falta de FT vem diminuindo de frequência de observação com o passar do amplo uso do ChatGPT pelo mundo, mas reforça a ideia de que uma sistematização como *Prompt*-EDU pode efetivamente impactar positivamente na otimização do *ChatGPT* quando aplicada por um tempo antes da requisição de tarefas complexas (Liu *et al*., 2023; Wang *et al*., 2023). Nesse cenário tem-se as respostas da conta IAT-BRA, classificadas como muito boas em dois níveis de verificação do processamento, mesmo na versão GPT3.5: [a] na LLM, mostrando que a CoT foi integrada com precisão na construção do FT específico e vinculado à taxonomia de Bloom (Rocha *et al*., 2021; Wang *et al*., 2023); [b] no NLP, quando generalizou corretamente a finalidade de um *prompt* educacional para docentes (Chen *et al*., 2023). Esses dois níveis do processamento reforçam a proposta de um roteiro como o *Prompt*-EDU e o teste na conta IAT-BRA mostrou que o roteiro tem potencial em otimizar o desempenho e melhorar a acurácia no tema ciências do processo ensino-aprendizagem (Liu *et al*., 2023).

Ainda é possível apontar um desempenho de FT satisfatório na conta IAT-BRA quando, na coleta de 29/04 (Quadro 7) o ChatGPT contextualizou corretamente o conceito de *prompt* educacional como um *"conjunto de instruções, perguntas ou sugestões que guiam o pensamento e a ação dos professores para alcançar um objetivo específico de ensino e aprendizagem, com base na Taxonomia de Bloom*". Efetivamente, vincular a significação dos verbos à taxonomia pareceu impactar sobre o viés das traduções não-controladas dentro do processamento, reduzindo problemas e erros de interpretação dos comandos potencialmente relacionados ao texto redigido no comando humano (Wang *et al*., 2023). Isso ficou mais evidente na declaração de finalidade gerada pela CoT, aplicada dentro da NLP: *"estruturar a comunicação com a inteligência artificial, fornecendo informações relevantes e específicas sobre o que se espera alcançar com a interação"*.

Além disso, a IA confirmou seu FT quando assumiu o protagonismo do seu personagem ao acatar o papel de usar *prompts* educacionais nos termos apresentados para *"direcionar o diálogo para o desenvolvimento de habilidades mais complexas, como a análise crítica, a síntese e a avaliação de informações, em vez de apenas fornecer respostas simples"*. Esses achados reforçaram os pontos observados na literatura: a necessidade de compreender o funcionamento do LLM para desenvolver um ML que otimize o desempenho do FT e desenvolvimento de um NLP específico (Liu *et al*., 2023) quando o objetivo é a aplicação no processo ensino-aprendizagem (Humble & Mozelius, 2022; Bearman *et al*., 2022; Crompton & Burke, 2023).

## Conclusões

Embora tenha mostrado uma otimização bastante satisfatória perante os resultados encontrados, o Demo-CoT é um *prompt* de preparo da máquina, dentro do roteiro *Prompt*-EDU e representa somente a primeira parte do FT e NLP na atribuição de tarefas mais complexas para fins educacionais. Ele é uma forma de desdobramento da tarefa complexa em estágios fragmentados, onde a natureza do prompt, e a

estruturação de sua redação com os elementos apropriados, facilita o domínio dos usuários menos experientes no desenvolvimento das especificidades requeridas em suas áreas e interfaces de atuação docente.

Esse estudo partiu do ponto mais amplamente discutido desde a chegada do ChatGPT ao público mundial: a forma de construir as melhores perguntas. Mas foi no seu desenho de desenvolvimento para a pesquisa que foi encontrado um ponto-chave inesperado: a inserção da taxonomia de Bloom como fator de controle das traduções converteu-se em um pilar central, ao invés de um apoio figurativo de controle do experimento. Aparentemente, vincular a taxonomia de Bloom pode ter sido uma chave de ajuste de desempenho na comunicação com a máquina, controlando vieses das traduções implícitas no processamento quando os *prompts* são realizados em idiomas não-nativos.

Essa experiência de vínculo da taxonomia em *prompts* para finalidades educacionais no ChatGPT é inédita nos estudos e tornou-se o elemento sobre o qual as novas pesquisas certamente se desdobrarão para o aprimoramento do roteiro *Prompt*-EDU. O caminho dos estudos com IA de texto, em especial com o ChatGPT, está apenas iniciando e há muito pela frente a ser explorado no âmbito das ciências de aprendizagem. O que se espera é que o roteiro e a sistematização *Prompt*-EDU seja um caminho para instrumentar docentes menos experientes em compreensão das redes neurais, para que possam trazer suas inestimáveis experiências de carreira e contribuir ativamente – como pesquisadores – no desenvolvimento de aplicações melhores e mais bem ajustadas da IA aos procedimentos e processos das ciências da aprendizagem.

## Referências bibliográficas

BARBOSA, Carlos Roberto de Almeida Correa. Transformações no ensino-aprendizagem com o uso da inteligência artificial: revisão sistemática da literatura. **RECIMA21-Revista Científica Multidisciplinar**, v. 4, n. 5, p. e453103-e453103, 2023. Disponível em: https://recima21.com.br/index.php/recima21/article/view/3103. Acesso em 10/05/2023.

Bearman, M., Ryan, J. & Ajjawi, R. Discourses of artificial intelligence in higher education: a critical literature review. *High Educ* (2022). Disponível em: https://doi.org/10.1007/s10734-022-00937-2. Acesso em 10/05/2023.

CELIK, Ismail *et al*. The promises and challenges of artificial intelligence for teachers: A systematic review of research. **TechTrends**, v. 66, n. 4, p. 616-630, 2022. Disponível em: https://link.springer.com/article/10.1007/s11528-022-00715-y. Acesso em 10/05/2023.

CHEN, Jiuhai *et al*. How Many Demonstrations Do You Need for In-context Learning? **arXiv**, n. 2303.08119v3, p. 1-11, 2023. Disponível em: https://arxiv.org/pdf/2303.08119.pdf. Acesso em 10/05/2023.

COTTON, Debby RE; COTTON, Peter A.; SHIPWAY, J. Reuben. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. **Innovations in Education and Teaching International**, p. 1-12, 2023. Disponível em: https://www.tandfonline.com/doi/full/10.1080/14703297.2023.2190148. Acesso em 10/05/2023.

CROMPTON, Helen; BURKE, Diane. Artificial intelligence in higher education: the state of the field. **International Journal of Educational Technology in Higher Education**, v. 20, n. 1, p. 1-22, 2023. Disponível em: https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00392-8. Acesso em 10/05/2023.

HUMBLE, Niklas; MOZELIUS, Peter. The threat, hype, and promise of artificial intelligence in education. **Discover Artificial Intelligence**, v. 2, n. 1, p. 22, 2022. Disponível em: https://link.springer.com/article/10.1007/s44163-022-00039-z. Acesso em 10/05/2023.

KASNECI, Enkelejda *et al*. ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences**, v. 103, p. 102274, 2023. Disponível em: https://www.sciencedirect.com/science/article/pii/S1041608023000195. Acesso em 10/05/2023.

LI, Yingcong *et al*. Transformers as Algorithms: Generalization and Stability in In-context Learning. **arXiv**, n. 2301.07067, p. 1-34, 2023. Disponível em: https://intra.ece.ucr.edu/~oymak/Transformers_as_algorithms.pdf. Acesso em 10/05/2023.

LIANG, Weixin *et al*. GPT detectors are biased against non-native English writers. **arXiv preprint arXiv:2304.02819**, 2023. Disponível em: https://arxiv.org/abs/2304.02819. Acesso em 10/05/2023.

LIAO, Wenxiong *et al*. Differentiate ChatGPT-generated and Human-written Medical Texts. **arXiv preprint arXiv:2304.11567**, 2023. Disponível em: https://arxiv.org/abs/2304.11567. Acesso em 10/05/2023.

LIU, Pengfei *et al*. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. **ACM Computing Surveys**, v. 55, n. 9, p. 1-35, 2023. Disponível em: https://dl.acm.org/doi/full/10.1145/3560815. Acesso em 10/05/2023.

MARZANO, Robert J.; KENDALL, John S. (Ed.). **The new taxonomy of educational objectives**. Corwin Press, 2006. Disponível em: http://dspace.vnbrims.org:13000/xmlui/bitstream/handle/123456789/4604/The-New-taxonomy-of-Educational-Objectives.pdf?sequence=1. Acesso em 10/05/2023.

DE ANDRADE MATIAS, Klesia; MORESI, Eduardo Amadeu Dutra; DOS SANTOS, Pricila Kohls. Tendências em inteligência artificial e educação híbrida: um estudo exploratório. **Poíesis Pedagógica**, v. 20, n. Publicação contínua, p. 76-96, 2022. Disponível em: https://periodicos.ufcat.edu.br/poiesis/article/view/73649. Acesso em 10/05/2023

OPEN AI INCORPORATED. GPT-4 Technical Report. **ArXiv**, n.2303.08774, 2023. Disponível em: https://arxiv.org/pdf/2303.08774.pdf. Acesso em 10/05/2023.

ROCHA, Daniyel NN; CAMPELO, Cláudio EC; JERÔNIMO, Caio LM. Aplicação de verbos como proxy para identificação automática do nível cognitivo de questões: uma abordagem baseada na taxonomia de Bloom. In: **Anais do XXXII Simpósio Brasileiro de Informática na Educação**. SBC, 2021. p. 897-908. Disponível em: https://sol.sbc.org.br/index.php/sbie/article/view/18116. Acesso em 10/05/2023.

SALAS-PILCO, Sdenka Zobeida; YANG, Yuqin. Artificial intelligence applications in Latin American higher education: a systematic review. **International Journal of Educational Technology in Higher Education**, v. 19, n. 1, p. 1-20, 2022. Disponível em: https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-022-00326-w. Acesso em 10/05/2023.

SHIN, Taylor *et al*. Autoprompt: Eliciting knowledge from language models with automatically generated prompts. **arXiv preprint arXiv:2010.15980**, 2020. Disponível em: https://arxiv.org/abs/2010.15980. Acesso em 10/05/2023.

VASWANI, Ashish *et al*. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html. Acesso em 10/05/2023.

WANG, Xinyi; ZHU, Wanrong; WANG, William Yang. Large language models are implicitly topic models: Explaining and finding good demonstrations for in-context learning. **arXiv preprint arXiv:2301.11916**, 2023. Disponível em: https://arxiv.org/abs/2301.11916. Acesso em 10/05/2023.

WEI, Jason *et al*. Chain of thought prompting elicits reasoning in large language models. **arXiv preprint arXiv:2201.11903**, 2022. Disponível em: https://arxiv.org/abs/2201.11903. Acesso em 10/05/2023.

ZAWACKI-RICHTER, Olaf *et al*. Systematic review of research on artificial intelligence applications in higher education–where are the educators? **International Journal of Educational Technology in Higher Education**, v. 16, n. 1, p. 1-27, 2019. Disponível em: https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-019-0171-0?fbclid=IwAR0vSk4s9y0V0vExpcAel6yL4LEb-PrNDnlreOB5WrGxlu8-3awpYGgK6Ig. Acesso em 10/05/2023.