

EVOLUÇÃO E O DESENVOLVIMENTO DO CHATBOT GPT-3 E GPT-4***EVOLUTION AND DEVELOPMENT OF CHATBOT GPT-3 AND GPT-4***

Jaqueline Francis Laureano Rosa - jaqueline.rosa01@fatec.sp.gov.br
Faculdade de Tecnologia de Taquaritinga (Fatec) – Taquaritinga – SP – Brasil

Diego Renan Bruno - diego.bruno01@fatec.sp.gov.br
Faculdade de Tecnologia de Taquaritinga (Fatec) – Taquaritinga – SP – Brasil

DOI: 10.31510/inf.v21i1.1849

Data de submissão: 14/03/2024

Data do aceite: 10/03/2024

Data da publicação: 20/06/2024

RESUMO

A *OpenIA* é uma empresa que vem se destacando no desenvolvimento da Inteligência Artificial para *chatbots* na atualidade, criando seu primeiro modelo chamado de GPT, que vem evoluindo a cada dia, trazendo novas versões como o GPT-3.5 e GPT-4. Essa evolução passou a utilizar técnicas de Inteligência Artificial que processavam somente texto para técnicas que processam imagens, revolucionando a arte de produzir vídeos, imagens, textos etc. Este artigo explora esta evolução e suas nuances, bem como, as técnicas de processamento demonstradas através de um estudo bibliográfico sobre o tema. Nas primeiras seções deste artigo foram expostos e explicados o que são os *chatbots* GPT. Já na sequência foi exposto um comparativo de desempenho entre os modelos GPT-3.5 e GPT-4, na outra seção foi explorado a arquitetura e os modelos com enfoque no aumento de seus parâmetros. Na seção seguinte foi avaliado o desempenho de cada modelo através de um exame do ENEM 2022 e por últimos, foi analisado as suas limitações e perigos desta ferramenta, como observadas os pontos que precisam melhorar e por fim, fizemos a conclusão destacando a notável evolução e a grande expectativa de avanços nos negócios e demais setores, destacando também a necessidade de tornar esta ferramenta um instrumento acessível para que todos possam participar destes avanços, tornando-a mais democrática. O objetivo deste trabalho foi analisar a evolução e as diferenças do *chatbot* GPT-3.5 para GPT-4 e os resultados demonstram um avanço em sua capacidade de entender a linguagem natural e compreender textos, como também processar imagens.

Palavra-chave: Inteligência Artificial. GPT-3.5. GPT-4.

ABSTRACT

OpenIA is a company that has stood out in the development of artificial intelligence for chatbots today, creating its first model called GPT, which has evolved every day, bringing new versions such as GPT-3.5 and GPT-4. This evolution has gone from using Artificial Intelligence techniques that only processed text to techniques that process images, revolutionizing the art of producing videos, images, text, etc. This article explores this evolution and its nuances, as well as the processing techniques demonstrated through a bibliographic study on the subject. The first sections of this article explain what GPT chatbots are. Next, a performance comparison

between the GPT-3.5 and GPT-4 models was presented, while another section explored the architecture and models with a focus on increasing their parameters. In the next section, the performance of each model was evaluated using an ENEM 2022 exam, and finally, the limitations and dangers of this tool were analyzed, as well as the points that need to be improved, and finally, we concluded by highlighting the remarkable evolution and the great expectation of advances in business and other sectors, also highlighting the need to make this tool an accessible tool so that everyone can participate in these advances, making it more democratic. The aim of this work was to analyze the evolution and difference of the GPT-3.5 to GPT-4 chatbot, and the results show an advance in its ability to understand natural language and text, as well as processing images.

Keywords: artificial intelligence. GPT-3.5. GPT-4.

1 INTRODUÇÃO

Nestes últimos tempos, houve avanços significativos no processamento de linguagem natural (PLN) e no aprendizado de máquina (AM) o que resultou em um alto desenvolvimento de Modelos de Linguagem (ML) sofisticadas. Estes modelos evidenciaram um notável dinamismo em diversas tarefas como; respostas a perguntas, tradução de textos, sumarização e muitas outras (NUNES et al, 2023).

A Inteligência Artificial (IA) surge com mais possibilidade e avanços ao longo da última década e se torna a tecnologia mais desejável em todo o mundo, trazendo consigo avanços em diversas áreas, possibilitando a pessoas leigas a capacidade de desenvolver novos usos em campos financeiros, comércio eletrônico etc.

Voltando na história, Alan Turing, considerado o pai da computação, mencionou o termo “Inteligência Artificial”, já em uma pesquisa da década de 1950. Conhecido como “jogo da imitação”, a investigação propunha um teste para avaliar se as máquinas possuíam capacidade de imitar o pensamento humano e se conseguiam passar por uma pessoa confundindo o humano do outro lado da tela. A partir do “Jogo da Imitação”, outras investigações foram feitas, aumentando o campo dessa pesquisa. Em 1957, surgiu o algoritmo Perceptron, primeira rede neural voltada para classificação de resultados. Dois anos depois, em 1959, o termo Machine Learning foi usado pela primeira vez, para explicitar que computadores haviam passado a ter a habilidade de aprender uma função mesmo sem serem programados diretamente para isso. Ou seja, os computadores passaram a saber como alimentar um algoritmo com dados de modo aprender a executar uma tarefa automaticamente. Em 1964, ELIZA, o primeiro chatbot do mundo, foi criado. ELIZA conversava imitando uma psicanalista e foi desenvolvida para usar respostas baseadas em palavras-chave e estrutura sintática¹⁷. Passaram-se mais de 70 anos entre as primeiras pesquisas e o surgimento de algoritmos capazes de interagir com humanos usando linguagem e diálogos – como o Chat GPT - mais refinados e baseados na IA generativa. E isso se deu porque só recentemente passou a existir poder computacional disponível e mais barato do que nos anos 1950/1960 para desenvolver a IA (SAMARÃO, 2023).

Este avanço considerável tem como ponto de partida o ChatGPT, uma ferramenta de inteligência artificial desenvolvida pela empresa OPENAI. Este modelo é baseado na linguagem generativa ou *Generative Pre-Trained Transformer* (GPT) com capacidade de responder e interagir com os usuários, ou seja, é um *chatbot* capaz de responder todo tipo de questão aos usuários.

Chatbot é um programa de computador que simula conversas entre a máquina e os humanos, ou seja, se trata de uma ferramenta projetada para interagir com o usuário respondendo todo tipo de questão, baseado em um modelo de aprendizado artificial, seja ele supervisionado ou não, para gerar uma linguagem natural semelhante a humana.

A popularização do *ChatGPT* foi muito rápida devido a sua capacidade de construir um simples texto até a mais complexos e com uma linguagem quase igual à do ser humano.

Apesar de toda essa desenvoltura, o *ChatGPT* tem suas limitações, pois foi construído por um enorme conjunto de dados de texto que podem possibilitar a difusão de preconceitos e estereótipos que porventura estejam inseridos em seus dados, podendo retornar ao usuário um texto preconceituoso ou ofensivo (SCIELO, 2023).

Shang (2023) escreve que o processamento de linguagem natural cresceu devido ao aparecimento de linguagens de grande dimensão que causaram um enorme impacto na tradução automática, na análise de sentimentos e sumarização de texto. O GPT-4 se destaca no desempenho destas funções, mas ainda traz problemas como raciocínio limitado, alucinação e vieses.

2 REFERENCIAL TEÓRICO

2.1 O que é *ChatGPT*

ChatGPT foi desenvolvido pela *OpenAi* como um modelo de linguagem de visualização. Seu treinamento é através da obtenção de uma enorme gama de variados textos da internet, possibilitando um vocabulário muito acurado e profundo conhecimento sobre variados temas. Quando ele recebe um texto como uma espécie de aviso, ele retorna uma resposta ao usuário de forma relevante e coerente.

Suas aplicações são muitas criativas, permitindo aos usuários um enorme acesso a diversas possibilidades, podendo ser explorado na produção de textos, composição de conteúdo e edição, gerar descrições e narrativas de peças visuais, na música pode gerar letras ou sugerir melodias, enfim, o *ChatGPT* é uma ferramenta incrível e versátil (MARQUES, 2023).

GPT significa *Generative Pre-trained Transformer*, um modelo de IA que usa redes neurais profundas para gerar linguagem natural a partir de um determinado prompt. A OpenAI desenvolveu essa poderosa tecnologia para gerar geração de texto legível por humanos a partir de vários conjuntos de dados. Usando métodos de aprendizagem auto-supervisionados, o GPT pode produzir resultados de linguagem natural altamente precisos em uma fração do tempo em comparação com as técnicas tradicionais de processamento de linguagem. A tecnologia generativa de transformadores pré-treinados oferece muitos benefícios para empresas e usuários individuais. As empresas podem usá-lo para tarefas como resumir o feedback do cliente ou criar recomendações de conteúdo com base em pesquisas de usuários. Também pode ajudar a agilizar processos como responder a consultas de clientes ou responder a avaliações on-line com rapidez e precisão – economizando tempo e recursos valiosos. Em um nível individual, os usuários podem se beneficiar de ter acesso a assistentes de IA mais brilhantes que podem entender melhor questões complexas sem exigir instruções de programação explícitas. A mídia social é outra área em que se pode usar modelos de linguagem generativa extensivamente. Usando o processamento de linguagem natural para análise de sentimentos, a OpenAI desenvolveu um algoritmo capaz de criar conversas significativas a partir de um prompt inicial. Em um contexto de mídia social, isso permite que os usuários criem conteúdo que envolva e forneça um toque mais personalizado (ANDRELUG, 2023).

2.1.1 Transformadores - o estado da arte no processamento da linguagem natural

Transformadores ou *Transformers* é um modelo de aprendizado profundo que teve seu início no ano de 2017 que tem um interessante mecanismo de atenção que pesa a influência das inúmeras partes dos dados de entrada. Este mecanismo é amplamente utilizado no Processamento de Linguagem Natural (PLN), que é um ramo da inteligência artificial que se concentra na interação entre computadores e seres humanos por meio da linguagem natural, mas seu desenvolvimento avançou e sua aplicação pode também realizar outras tarefas, como por exemplo: compreensão de vídeo.

Estes transformadores foram projetados para o uso de resumo e tradução de texto através de dados de entrada sequenciais, utilizando pela linguagem natural, assim como as Redes Neurais Recorrentes (RNN), que são redes neurais usadas na tradução de idiomas, reconhecimento de fala e legendagem de imagem, com a vantagem de não necessitar que estes dados sequenciais sejam processados em ordem, mas usa a operação de atenção para identificar o contexto de qualquer posição na sequência de entrada (DATA SCIENCE ACADEMY, 2023).

Nesta situação se os dados de entrada forem de uma frase em linguagem natural, o Transformador ao invés de processar os dados de início para o final, ele identifica o contexto que traz o significado a uma determinada palavra que está contida na frase, permitindo assim mais paralelização do que as RNNs, diminuindo significativamente o tempo de treinamento.

Devido a isso, os transformadores se tornaram o principal modelo para a resolução de problemas de PLN, substituindo os modelos de redes neurais recorrentes. Esta escolha se deve, principalmente, pela facilidade de paralelização que o transformador permite durante o treinamento, tornando possível o treinamento de um maior conjunto de dados, resultando na criação de sistemas pré-treinados como *Generative Pre-Training Transformer* (GPT), que são treinados com grandes quantidades de dados de linguagem geral e podem ter suas tarefas ajustadas para tarefas específicas de linguagem (DATA SCIENCE ACADEMY, 2023).

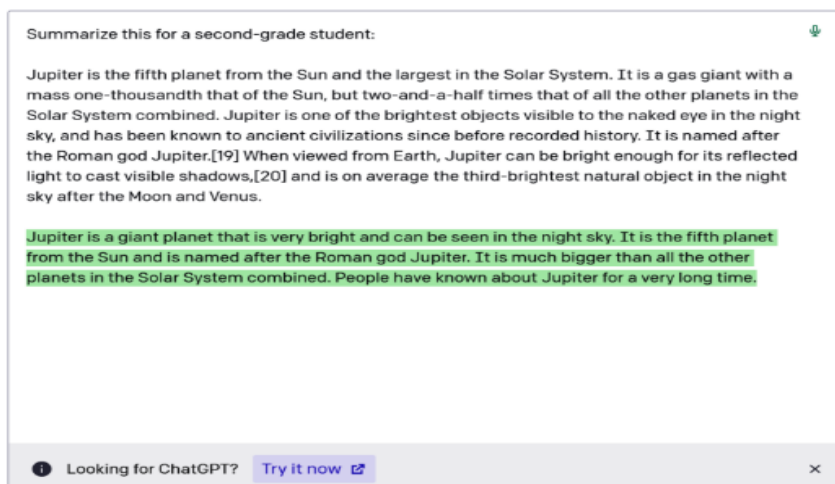
2.2 GPT 3

Apesar de serem os precursores do GPT-3, os modelos GPT-1 e GPT-2 foram muito limitados, apesar da significativa conquista no processamento de linguagem natural. O modelo GPT-1 era propenso a gerar textos repetitivos e falhava no rastreamento de dependências de longo prazo no texto, enquanto o GPT-2 falhava na compreensão de tarefas que exigiam raciocínio mais complexo, porém, estas limitações abriram espaço para o desenvolvimento do modelo GPT-3.

GPT-3 é a terceira versão que foi aprimorada pela OpenAI no ano de 2018, atualizando e sobrepondo a versão GPT-2, pois foi treinada com uma base maior de dados. Esta versão vem com uma precisão maior, pois tem até 175 bilhões de parâmetros, fornecendo assim, uma maior precisão na elaboração de textos, ainda que tenha um tamanho maior, seu modelo exige menos dados que outros modelos de linguagem como BERT e XLNET para atingir resultados satisfatórios. Ele se mostrou mais eficiente ao realizar tarefas de processamento de linguagem natural como também a resultados sobre respostas a perguntas. Este *chatbot* é uma ferramenta de ciência de dados que pode realizar inúmeras tarefas, como resumir o feedback do cliente ou fazer recomendações de conteúdo baseado na pesquisa do usuário (ANDRELUG, 2023).

O modelo GPT-3 pode entender e gerar linguagem natural, pois foi treinado com uma grande camada de corpos de texto e isto lhe capacitou a executar uma vasta gama de tarefas em linguagem natural, como resumo de textos, tradução de idiomas, perguntas e respostas e muito mais (ALTO, 2023). Na figura a seguir está demonstrado um resumo de um texto usando o *ChatGPT*.

Figura 1: Exemplo de um resumo usando o GPT-3



Fonte: ALTO, 2023.

2.3 GPT 3.5

O *ChatGPT 3.5* é um modelo construído com base no *InstructGPT* que é da série de modelos GPT 3.5. Ele foi desenvolvido como um modelo que adapta um corpus de ajuste de instrução em larga escala para *InstructGPT* em uma configuração conversacional.

O modelo GPT-3 foi atualizado com o uso da aprendizagem por reforço através do *feedback* humano, este modelo utiliza o padrão de recompensas para obter melhor qualidade na geração aproximando com a geração humana.

Um dos maiores benefícios do modelo *InstructGPT* é o seu avanço na capacidade de obter melhores resultados com as instruções de língua inglesa comparado ao modelo GPT-3. Há três modelos principais de GPT-3.5 que são: *code-davinci-002*, *text-davinci-002* e *InstructGPT*. Estas séries foram projetadas para criar conteúdo seguros com a finalidade de criar falsos textos, textos tóxicos ou prejudiciais (ESPEJEL et al, 2023).

O modelo GPT-3.5 é baseado em *text-davinci-002* e por isso é considerado mais avançado e poderoso que os modelos anteriores, pois tem a capacidade de gerar linguagem e acompanhamento de instruções de nível superior.

2.4 GPT 4

Este é o modelo que superou todos os anteriores, com uma versão multimodal poderosa e capacidade de processar, além de entradas textuais, também entradas visuais. Seu treinamento

usa uma incomparável quantidade de dados e grande poder computacional, demonstrando melhorias significativas que seus modelos anteriores (ESPEJEL et al, 2023).

O GPT-4 é um modelo multimodal que tem a capacidade de processar as entradas de imagem e texto e retornar como saída um texto. Desta forma, este *chatbot* é um modelo que tem um enorme potencial de uso em diversas aplicações, como resumo e tradução de texto, além de um sistema de diálogo.

Desenvolvido para melhorar a compreensão e gerar texto em linguagem natural, o GPT-4 pode enfrentar cenários complexos e de muitas nuances, pois foi avaliado em uma variedade de exames que foram desenvolvidos para humanos. Nestas avaliações seu desempenho demonstrou uma grande evolução, pois em certos testes ele superou até os seres humanos (OPEN AI, 2023).

O GPT-4 e o GPT-3.5 compartilham vários recursos em comum, pois os dois modelos têm a arquitetura de transformadores, porém com escalas diferentes. No modelo transformador há uma dependência de arquitetura codificador-decodificador com módulos de auto atenção que capturam relacionamentos complexos e conseguem extrair padrões de sequências de entrada. Neste sentido, o codificador faz o processamento das sequências de entrada, enquanto o decodificador converte a saída do codificador gerando uma sequência na saída do transformador.

Apesar de a OpenAI não revelar muitas informações sobre o processo de treinamento do GPT-4, fica claro que a abordagem deste modelo é a de recompensa baseada em regras (RBRM) quando comparada com seu antecessor. Além deste aprendizado, o GPT-4 também é treinado com o Aprendizado por Reforço com *FeedBack* Humano (RBRMs). Esta abordagem de modelos baseada em regras (RBRMs) melhora significativamente o seu desempenho e a segurança, gerando sinais de recompensa adicionais no momento em que o processo de ajuste fino do Aprendizado por Reforço com Feedback Humano sobre o texto gerado acontece, garantindo a harmonia com a produção de conteúdo seguro e correto (KOUBAA, 2023).

3 PROCEDIMENTOS METODOLÓGICOS

A metodologia utilizada para o desenvolvimento deste trabalho foi a bibliográfica e foi explorado sites, livros e trabalhos acadêmicos como fonte de pesquisa. Na pesquisa bibliográfica é fundamental que o pesquisador levante dados a respeito do tema abordado e consulte trabalhos que foram desenvolvidos por outros pesquisadores, pois assim, os conceitos

e aspectos referente a pesquisa podem ser explorados e conhecidos (LAKATOS; MARKONI, 2003).

4 RESULTADOS E DISCUSSÃO

Um estudo recente aplicou um exame da Ordem dos Advogados dos Estados Unidos ao GPT-3.5 e obteve a aprovação em duas categorias dentre sete testadas. Outro estudo que fez testes de raciocínio quantitativo analítico dos exames de Contadores Públicos Certificados (CPA), demonstrou um resultado ruim do GPT-3.5 quando comparado com o desempenho humano, mas demonstra uma melhor capacidade quando se trata de questões que exigem habilidades de entender, lembrar e aplicar conhecimento, atingindo o nível de entendimento comparável ao humano (NUNES et al, 2023).

Outro experimento realizado utilizando *Large Language Model (LLM)* ou Grandes Modelos de linguagem em eventos históricos e resultados potenciais tiveram sua avaliação em uma métrica chamada Distancia a Realidade (DTR) que quantifica a disparidade entre a realidade histórica e as previsões da inteligência artificial. Dentro desta métrica, uma pontuação mais baixa indica que a IA está mais perto da realidade ou seja, é mais precisa em sua previsão.

O teste foi realizado entre GPT-3.5, GPT-3.4 como também o BARD (Bard do Google é um *chatbot* de inteligência artificial que compete com o ChatGPT da *OpenAI*), conforme segue a tabela abaixo;

Tabela 1 – Comparativo de desempenho entre Inteligências Artificiais

TABELA 1	Distância da realidade		
Evento nº	GPT-3.5	GPT-4	Google-BARD
1	0,1	0,1	0,1
2	0,2	0,05	0,3
3	0,1	0,05	0
4	0,3	0,01	0,6
5	0,1	0	0,2
6	0,1	0	0
Média	0.15	0.035	0.20

Fonte: TAŞAR e TAŞAR, 2023.

Atentando a diferença entre o GPT-3.5 E GPT-4 que é o objetivo deste estudo, a tabela acima demonstra que a pontuação média do GPT-4 foi mais baixa que as demais, ou seja, este desempenho superior se deve ao conjunto de dados de treinamento maior e a arquitetura avançada que permite ao GPT-4 melhor compreensão dos contextos históricos mais complexos(TAŞAR e TAŞAR, 2023).

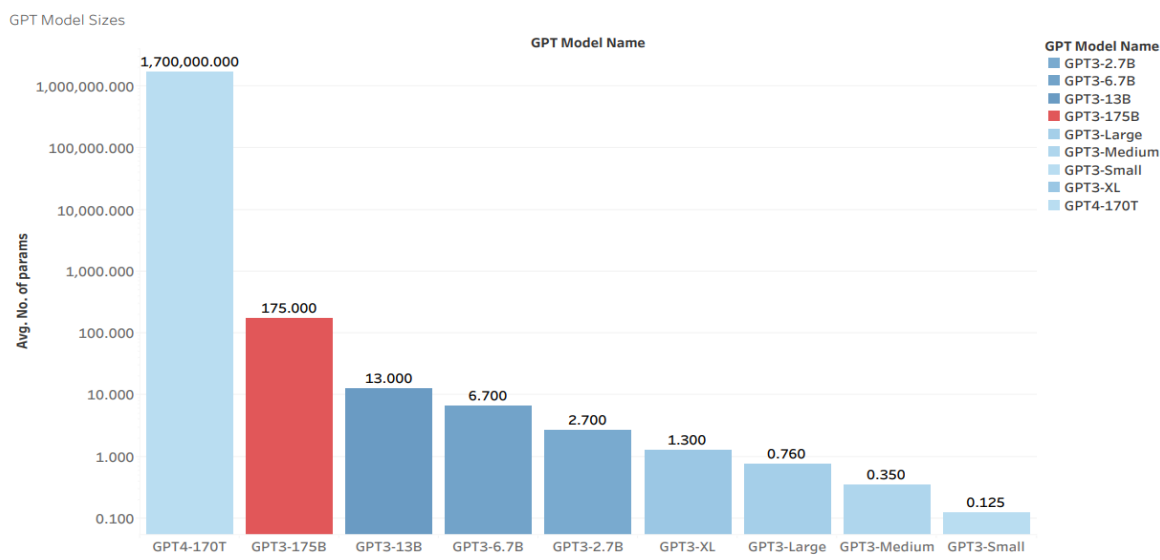
4.1 Arquitetura e modelo

O GPT-4 tem muitos recursos comuns com o GPT-3.5, pois os dois tem uma dependência do modelo arquitetônico transformador, mas em escalas distintas. Este modelo transformador depende de uma arquitetura codificador – decodificador que possuem módulos de auto atenção que buscam relacionamento complexos onde são extraídos os padrões de sequencias de entrada. Neste esquema o codificador processa a sequência de entrada e na saída do transformador, o decodificador converte a saída do codificador.

Apesar de a *OpenIA* não divulgar detalhes sobre a arquitetura do GPT-4, ela afirma que ele é um modelo estilo Transformador pré-treinado para prever o próximo token em um documento que são acessados em dados fornecidos pela internet ou por fornecedores terceirizados e então ajustados para usar Aprendizado por Reforço a partir do *feedback* humano (KOUBAA, 2023).

A arquitetura do GPT-4 avançou significativamente em escala quando comparamos com as IAs anteriores. A figura a seguir demonstra as diferenças entre os modelos GPT, com enfoque no expressivo aumento nos parâmetros presente no GPT-4.

Figura 2: Comparação entre modelos GPT



Fonte: Kobaa, 2023.

O modelo GPT-4 é mil vezes maior que o GPT-3.5, pois atinge 170 trilhões de parâmetros enquanto o GPT-3.5 atinge 175 bilhões. Esta diferença demonstra a capacidade ampliada do GPT-4 que se converte em uma precisão e desempenho, capaz de lidar com

modelos de linguagem complexas e com tarefas de processamento de linguagem natural, além de possibilitar a entrada de dados multimodais, incluindo imagens e textos (KOUBAA, 2023).

4.2 Algumas avaliações de desempenho

O teste em exames acadêmicos demonstra um avanço do GPT-4 sobre o seu antecessor, pois ele supera o GPT-3.5 em vários destes exames. Um teste no *Uniform Bar Exam* dos Estados Unidos, que é a prova semelhante ao exame da OAB no Brasil, revelou a impressionante capacidade jurídica do GPT-4, demonstrando seu avanço no processamento de linguagem natural e se equiparando ao desempenho humano nestes tipos de exames (KOUBAA, 2023).

Nunes Et al (2023) avaliou o exame do ENEM 2022 utilizando as questões e alternativas de várias edições do exame do ENEM. Neste trabalho os pesquisadores identificaram as questões marcando cada uma com um domínio, como: compreensão de texto, conhecimento enciclopédico e específico do domínio, elementos químicos e compreensão de imagens. Estes domínios foram essenciais para determinar e identificar se o elemento é diferente de um texto, como por exemplo, uma imagem ou símbolos químicos.

Para a avaliação do desempenho foi usado o conjunto de dados do ENEM 2022 com questões que se tornaram públicas após o treinamento dos modelos, dados estes que foram essenciais para a comparação dos modelos GPT-3.5 e GPT-4.

O resultado foi que o GPT-4 teve uma acurácia de 87% no exame de 2022, enquanto o modelo GPT-3.5 teve um desempenho de 76%. Este resultado levou os autores a concluir que ambos os modelos tem capacidade de resolver questões complexas do exame do ENEM e este avanço tem implicações importantes a serem consideradas no âmbito da educação e possibilita o desenvolvimento de novos modelos de linguagem no futuro (NUNES et al, 2024).

Em testes que teve como foco a língua inglesa, o desempenho do GPT-4 apresentou uma melhora significativa em relação ao GPT-3.5 e em testes com outras línguas o desempenho foi superior a 80% (KOUBAA, 2023).

4.3 Limitações do GPT-4

Apesar dos grandes avanços do GPT-4 este modelo tem algumas limitações que são igualmente encontradas no GPT-3.5, entretanto, seu impacto é reduzido.

Dentre as limitações, podemos destacar:

- Alucinação: este é um dos problemas mais sérios que são encontrados nos modelos da *OpenAI*. Estas alucinações são quando a inteligência artificial

produz conteúdo sem sentido ou impreciso, apesar de uma considerável melhora do GPT-4 em relação ao GPT-3.5.

- Discursos de ódio: Todos os modelos de IA tem potencial de produzir conteúdo nocivos, como discurso de ódio e incitação à violência. O GPT-4 apresentou um melhor desempenho com relação a estas questões.
- Operação de Desinformação e influência: neste caso a IA influencia a pessoa agir de forma errada e isso ocorre porque a IA busca suas informações em diferentes canais de comunicação e isso pode levar a desinformação. Na comparação com o GPT-3.5, o GPT-4 é melhor na mitigação destes conteúdos, mas ainda pode ser usado para este fim.

5 CONCLUSÃO

Neste estudo foi analisado o avanço do *chatbot* modelo GPT e na comparação de seus modelos, conseguimos identificar uma melhora significativa no modelo GPT-4.

Entre os avanços encontrados, foi na sua capacidade de compreensão histórica da linguagem humana, superando seu antecessor GPT-3.5, principalmente por permitir a entrada de imagens como dado, resultado do aumento de seu tamanho, atingindo 170 trilhões de parâmetros processados. Este tamanho possibilita ao GPT-4 um alto desempenho e uma ótima precisão, bem como, maior capacidade para lidar com modelos de linguagem complexos e com linguagem natural.

Outro destaque observado é o salto em multimodalidade, pois o GPT-4 foi treinado com grande quantidade de dados e poder computacional que resultou em melhorias significativas na compreensão de textos e superou desafios complexos, com exames projetados para humanos, além de ter aprimorada a capacidade de seguir instruções e gerar resposta mais alinhadas com as solicitações.

REFERÊNCIAS

ANDRELUG. GPT-4 vs GPT-3 – A comparação completa. Disponível em: <https://andrelug.com/gpt-4-vs-gpt-3/>. Acesso em: 06 fev. 2024.

CHANG, Edward. *Examining GPT-4's Capabilities and Enhancement with SocraSynth*. 2023. Disponível em: https://www.researchgate.net/publication/374753069_Examining_GPT-4%27s_Capabilities_and_Enhancement_with_SocraSynth. Acesso em 04 fev. 2023.

DATA SCIENCE ACADEMY. Deep Learning Book. Redes Neurais, 2022. Disponível em: <https://www.deeplearningbook.com.br/?s=redes+neurais>. Acesso em: 17 Setembro 2023.

ESPEJEL, J.L.; ETTIFOURI, EL H.; ALASSAN, M.S.Y.; CHOUHAM, EL M.; DAHHANE, W. *GPT-3.5, GPT-4, or BARD? Evaluating LLMs reasoning ability in zero-shot setting and performance boosting through prompts*. *Natural Language Processing Journal*, v.5, 2023. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2949719123000298>. Acesso em: 06 fev. 2023.

KOUBAA, A. GPT-4 vs. GPT-3.5: um confronto conciso. TechRxiv. Pré-impressão.. Disponível em: <https://www.techrxiv.org/doi/full/10.36227/techrxiv.22312330.v1>. Acesso em: 10 fev. 2024.

MARCONI, Marina Andrade., LAKATOS, Eva Maria. Fundamentos de metodologia científica. 5º ed. São Paulo: Atlas, 2003.

MARQUES, J.F. **Bíblia do Chat Gpt–10 livros em 1**. Clube de Autores. 2023. Disponível em: https://www.google.com.br/books/edition/B%C3%ADblia_Chat_Gpt_10_Livros_Em_1/m7nsEAAAQBAJ?hl=pt-BR. Acesso em: 06 fev.2024

NUNES, D., Primi, R., PIRES, R., LOTUFO, R.D., & NOGUEIRA, R. *Evaluating GPT-3.5 and GPT-4 Models on Brazilian University Admission Exams*. ArXiv, abs/2303.17003. 2023. Disponível em: <https://arxiv.org/abs/2303.17003>. Acesso em 05 fev. 2024.

OPEN AI. *GPT-4 Technical Report*. Disponível em: <https://cdn.openai.com/papers/gpt-4.pdf>. 2023. Acesso em 08 fev. 2024.

SAMARÃO, L. A. **IA e algoritmos enviesados**: as dimensões do viver na sociedade 5.0. Tese. Universidade Católica de São Paulo. 2023.

SCIELO. ChatGPT: um museu de grandes novidades. **Cadernos EBAPE.BR**, v. 21, n. 1. 2023. Disponível em: <https://www.scielo.br/j/cebape/a/FHBLtCcQndXVLGSZhQqnmWn/#>. Acesso em 05 fev. 2024.

SINGH, SHIVENDRA PRATAP. “Comparing GPT-4 and GPT-3.5: Advancements and Key Differences.” Towardsprompt, 2023.

Tasar, Davut Emre, and Ceren Ocal Tasar. "Bridging History with AI A Comparative Evaluation of GPT 3.5, GPT4, and GoogleBARD in Predictive Accuracy and Fact Checking." arXiv preprint arXiv:2305.07868. 2023. Disponível em: <https://ui.adsabs.harvard.edu/abs/2023arXiv230507868T/abstract>. Acesso em 03 Fev. 2023.

VALENTINA, A. *Modern Generative AI with ChatGpt and OpenAi models*. Packt Publishing Lts. Birmingham, UK. 2023.