



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Um estudo sobre o ChatGPT aplicado ao ensino e aprendizagem: benefícios, riscos e recomendações

Plínio Candide Rios Mayer
Marco Antônio Nemetala Garcia

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientador
Prof. Dr. Jan Mendonça Corrêa

Brasília
2025

Dedicatória

Plínio Candide Rios Mayer

Eu dedico esse trabalho à minha mãe, falecida em 24/09/2023, durante a confecção desse trabalho. Que, de onde ela estiver, ela possa ver seu filho se formando.

Marco Antônio Nemetala Garcia

Dedico a todos que, na tecnologia, encontram um caminho para o saber e o esclarecimento.

Agradecimentos

Plínio Candide Rios Mayer

Eu gostaria de agradecer, em primeiro lugar, ao nosso professor, Jan Corrêa, sem o qual esse trabalho não seria possível. Um orientador que adaptou a defesa deste trabalho para um prazo razoável para mim e para o Marco Antônio e sempre se mostrou solícito e colaborativo quanto às nossas demandas. Em segundo lugar, a um amigo cuja amizade sobreviveu à ardua tarefa de confecção de um trabalho de conclusão de curso, meu parceiro, Marco Antônio. Gostaria de agradecer, também, à minha irmã, minha única parente consanguínea próxima durante a confecção deste trabalho, além de companheira pela vida inteira. Por fim, aos meus amigos: Anderson Júnior, Lucas Bajo, Liliana Georg e Gabriel Almeida, que não desistiram de mim e salvaram o pouco que restava da minha sanidade.

Marco Antônio Nemetala Garcia

Gostaria de começar agradecendo ao professor Jan, pela valiosa orientação que permitiu desenvolver esse trabalho. Sou muito grato ao Plínio por embarcar junto comigo nesta jornada, me acompanhando não só na elaboração desse trabalho, mas ao longo do curso inteiro. Essa jornada foi bem mais fácil e divertida com você ao meu lado. Agradeço a minha família e amigos por me apoiarem dentro e fora da faculdade e por me incentivarem até o fim. Agradeço a todos os professores que cruzei por me ajudarem a formar meu senso crítico e a construir uma base sólida para minha formação acadêmica. Agradeço a cada um por compartilhar seus conhecimentos, experiências e paixão pelo ensino, elementos que moldaram/moldam o cientista que estou me tornando.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Acesso ao Portal de Periódicos.

Resumo

Este trabalho consiste em uma análise bibliográfica sobre as consequências da introdução de *LLMs* no processo ensino-aprendizagem, em especial do *ChatGPT*. Buscou-se obter informações sobre a utilização atual de ferramentas baseadas em inteligência artificial no processo ensino-aprendizagem, bem como possíveis aplicações e dilemas éticos relacionados ao seu uso. O estudo explora conceitos associados ao desenvolvimento de inteligências artificiais, desde os conceitos mais básicos como inteligência artificial, aprendizado de máquina, aprendizado profundo e viés algorítmico, até conceitos mais complexos como a arquitetura *transformer*, utilizada pelo *ChatGPT*. Em relação ao *GPT*, modelo por trás do *ChatGPT*, foram reunidos dados sobre seu histórico, desenvolvimento e evolução. Foram reunidos, também, dados relacionados às capacidades e limitações da ferramenta em várias áreas do conhecimento como direito e medicina, embora tenha-se dado foco às capacidades relacionadas à matemática e programação. Para facilitar a compreensão, buscou-se organizar os fenômenos ligados à introdução da inteligência artificial no processo ensino-aprendizagem em benefícios, riscos e recomendações. O ponto central dessa abordagem é a representação em forma de tabela, na qual são apresentados os riscos e as recomendações propostas por cada autor, proporcionando uma visão estruturada do tema no contexto do ensino superior.

Palavras-chave: inteligência artificial, ChatGPT, ensino superior, ética, educação, LLM, IA generativa

Abstract

This paper consists of a bibliographical analysis on the consequences of introduction of *LLMs* in the teaching-learning process, especially *ChatGPT*. It seeks to obtain information on the current use of tools based on artificial intelligence in the process teaching-learning, as well as possible applications and ethical dilemmas related to its use. This study explores concepts associated with development of artificial intelligence, from the most basic concepts such as artificial intelligence, machine learning, deep learning and algorithmic bias, even more complex concepts such as the *transformer* architecture used by *ChatGPT*. Regarding the *GPT*, the model behind *ChatGPT*, data was gathered about its history, development and evolution. Data were also gathered related to the capabilities and limitations of the tool in various areas of the knowledge such as law and medicine, although focus has been placed on mathematics and programming. To facilitate understanding, it organizes the phenomena linked to the introduction of intelligence artificial in the teaching-learning process in benefits, risks and recommendations. The central point of this approach is the representation in table form, which presented the risks and recommendations proposed by each author, providing a structured view of the topic in the context of higher education.

Keywords: artificial intelligence, ChatGPT, higher education, ethics, education, LLM, generative AI

Sumário

1	Introdução	1
1.1	Justificativa	1
1.2	Problema	2
1.3	Objeto	3
1.4	Objetivo Geral	3
1.5	Objetivos Específicos	3
1.6	Estrutura do Trabalho	4
2	Metodologia	5
2.1	Limitações	5
3	ChatGPT	7
3.1	Inteligência Artificial	9
3.2	Aprendizado de Máquina	10
3.3	Viés Algorítmico	14
3.4	Aprendizado Profundo	14
3.5	<i>LLMs</i>	17
4	GPT	20
4.1	História	20
4.2	Pré-treinado (<i>Pre-trained</i>)	21
4.3	<i>Transformer</i>	21
4.3.1	Codificador (<i>Encoder</i>)	22
4.3.2	Decodificador (<i>Decoder</i>)	22
4.3.3	<i>Embeddings</i>	23
4.3.4	Mecanismos de Atenção (<i>Attention Mechanisms</i>)	24
4.4	Tokenização	24
4.5	Capacidades e limitações	26

5	Os <i>LLMs</i> (<i>ChatGPT</i>) no processo ensino-aprendizagem	30
5.1	Inteligência artificial vs. inteligência humana	33
5.2	Aprendizado com <i>ChatGPT</i>	33
6	Benefícios dos <i>LLMs</i> (<i>ChatGPT</i>) no processo ensino-aprendizagem	36
6.1	Criação de avaliações de aprendizagem	37
6.2	Aprimoramento da prática pedagógica	38
6.3	Oferecimento de tutoria personalizada	40
6.4	Criação de esboços de artigos e pesquisas	41
6.5	Debate de ideias	42
6.6	Conscientização ética	43
6.7	Acesso à informação	43
6.8	Redução da ansiedade	44
7	Riscos dos <i>LLMs</i> (<i>ChatGPT</i>) no processo ensino-aprendizagem	45
7.1	Riscos à integridade acadêmica	45
7.2	Avaliações de aprendizagem injustas	46
7.3	Informação imprecisa	47
7.4	Perpetuação de preconceitos	47
7.5	Dependência excessiva da IA	48
7.6	Riscos à privacidade	49
8	Recomendações dos <i>LLMs</i> (<i>ChatGPT</i>) no processo ensino-aprendizagem	50
8.1	Riscos à integridade acadêmica	50
8.2	Avaliações de aprendizado injustas	51
8.3	Informação imprecisa	51
8.4	Dependência excessiva da IA	52
8.5	Perpetuação de preconceitos	52
8.6	Riscos à privacidade	52
9	Resultados	54
10	Conclusões	64
11	Trabalhos Futuros	67
	Referências	69

Lista de Figuras

3.1	Resposta do ChatGPT para a pergunta "o que é ChatGPT". Observe que além de responder, o ChatGPT ainda gera um título para a interação: "ChatGPT: Modelo PLN Poderoso".	8
3.2	Relação entre os conceitos utilizados no campo da inteligência artificial . .	9
3.3	Visualização dos conjuntos de documentos baseado no exemplo apresentado em <i>Foundations of Machine Learning</i> [31].	11
3.4	Abstração da organização dos arquivos dentro do algoritmo após treinamento. a linha pontilhada é o resultado do balanceamento dos parâmetros.	12
3.5	Exemplo de codificação da publicação <i>Computing Machinery and Intelligence</i> de Alan Turing no formato <i>BibText</i>	13
3.6	Imagem simulando um <i>captcha</i> , ferramenta usada para identificar se um usuário é humano. A ideia é justamente que é difícil para um computador conseguir compreender que imagens contém a representação de um carro. .	15
3.7	Representação de uma Perceptron Multicamadas (<i>MLP</i> em um grafo ordenado. Cada vértice do grafo representa um perceptron, enquanto as arestas representam as entradas e saídas de cada perceptron.	16
3.8	Imagem representando um modelo de aprendizado profundo onde cada coluna representa uma das camadas de representação da informação.	17
3.9	<i>LLMs</i> como o resultado do treinamento de algoritmos de aprendizado profundo utilizando grandes quantidades de dados.	18
4.1	Modelo <i>Transformer</i> segundo Vaswani[53]	25

4.2 Gráfico obtido de [22]. Tradução livre da legenda original: "Classificação média para cada arquivo em cada subdataset de GHOSTS em 9 de janeiro de 2023 e as versões de 30 de janeiro de 2023 do *ChatGPT* e para miniGHOSTS no GPT-4. Observe que o máximo a classificação é 5 e a classificação mínima, onde a questão foi pelo menos compreendida, é 2, ver Apêndice B.4; a classificação mais baixa de 1 indica que a resposta erra completamente a pergunta. Assim, uma passagem razoável a nota, ou seja, 50% dos pontos, corresponde à nota 3,5, indicada pela linha pontilhada vertical. O barras de erro representam intervalos de confiança de 95%." [22] 27

Lista de Tabelas

4.1	Tabela retirada do relatório publicado pela OpenAI [12]. Os resultados expressos em porcentagem representam a porcentagem de participantes abaixo do GPT, ou seja, um resultado de 90% indica que o GPT se saiu melhor do que 90% dos participantes.	29
9.1	Tabela contendo o resumo dos benefícios e riscos apontados pelos autores das publicações estudadas.	55
9.2	Tabela contendo os benefícios e riscos segundo cada um dos autores estudados para este trabalho	56
9.3	Tabela contendo o resumo das recomendações feitas pelos autores para os riscos apontados pelas publicações estudadas.	63

Capítulo 1

Introdução

1.1 Justificativa

Segundo [1] e [2], os últimos anos foram marcados pelo desenvolvimento e adoção de modelos de inteligência artificial em várias áreas da sociedade. Essas ferramentas, baseadas em inteligência artificial, têm sido amplamente utilizadas no ambiente acadêmico. O *ChatGPT*, um modelo de linguagem treinado em uma grande quantidade de dados textuais, para gerar respostas coerentes e contextualmente relevantes, é um exemplo dessas ferramentas. Seguindo o mesmo raciocínio, os estudos de [2], [3] e [4] consideram que a introdução desses modelos, apoiada pela ubiquidade de aparelhos digitais portáteis, como *smartphones* e *tablets* na vida cotidiana de muitas pessoas, ocasionou uma profunda mudança na forma como as pessoas interagem com a tecnologia.

De acordo com [5], o *ChatGPT* não se destaca apenas por ter facilitado a interação de pessoas comuns com modelos de inteligência artificial. O seu lançamento marca o uso da inteligência artificial não somente para compreensão da informação, mas também para a geração de novas informações, demonstrando possuir imenso valor para várias áreas da indústria, como cultura, educação e medicina. De acordo com [6], isso é possível graças à imensa capacidade de produzir grandes quantidades de dados a um custo relativamente baixo.

Além disso, é de entendimento de [4] que o uso de ferramentas baseadas em inteligência artificial tem se mostrado cada vez mais importante e relevante dentro do ambiente acadêmico. Essas ferramentas podem auxiliar tanto os estudantes quanto os professores em diversas atividades, como a realização de pesquisas, a busca por informações e o desenvolvimento de projetos. Através da inteligência artificial, é possível obter respostas rápidas e precisas para dúvidas acadêmicas, facilitando a consulta de informações e ampliando as possibilidades de acesso ao conhecimento.

1.2 Problema

É por conta dessa utilização que, assim como trazido por [7], faz-se necessário atentar-se aos potenciais prejuízos que esse tipo de ferramenta pode causar no processo de ensino-aprendizagem. Dentre eles, vale citar: o potencial para facilitar a fraude acadêmica; a geração de informações incorretas ou falsas; a capacidade de burlar detectores de plágio; a limitação do conhecimento; a falta de atualização com dados recentes; e a possibilidade de comprometer a integridade acadêmica. Além disso, há preocupações sobre a capacidade do *ChatGPT* de gerar conteúdo original e a possibilidade de que seu uso possa prejudicar a avaliação justa do desempenho dos alunos.

Segundo [4], a preocupação causada principalmente pelo medo de plágio e da terceirização do trabalho por parte dos estudantes levou a proibição do uso do *ChatGPT* por parte de várias instituições ao redor do mundo. Os desafios enfrentados pelos professores ao lidar com o uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico são diversos. De acordo com [8], um dos principais desafios é a necessidade de atualização constante e adaptação às novas tecnologias. Os professores precisam estar familiarizados com as funcionalidades e limitações dessas ferramentas para utilizá-las de forma eficaz em suas práticas pedagógicas.

Essas são apenas algumas das preocupações apresentadas pelos docentes, que, segundo [9], apresentam resistência à utilização de ferramentas baseadas em inteligência artificial. Dentre os motivos para essa resistência, estão, ainda de acordo com [9], a dificuldade de adaptação a novas tecnologias e o medo da substituição por essas ferramentas automatizadas.

Os discentes também são afetados pela introdução desse tipo de ferramenta. Segundo [10], a falta de interação humana pode ser prejudicial para os estudantes, uma vez que a máquina não é capaz de compreender plenamente as nuances das perguntas e pode fornecer respostas com conteúdo superficial. Ademais, citando [10] novamente, há, ainda, a possibilidade de perda da criatividade e originalidade nos trabalhos acadêmicos, uma vez que os estudantes podem se limitar a reproduzir as respostas geradas pela ferramenta.

Ainda assim, a pesquisa disponível em [11] considera que tanto os alunos quanto as faculdades geralmente possuem um olhar positivo em relação ao *ChatGPT*, e os estudantes não o consideram uma ameaça, desde que as informações fornecidas por ele sejam verificáveis. Ademais, é de entendimento de [11] que, se as escolas fugirem da adoção de tecnologias emergentes, isso poderá criar uma dicotomia entre as tecnologias utilizadas pelas instituições de ensino e aquelas usadas pelos alunos.

1.3 Objeto

De acordo com [6], o *ChatGPT* alcançou mais de 100 milhões de usuários logo após seu lançamento, tornando-o uma das ferramentas de inteligência artificial mais utilizadas pelo público. Por essa razão, ele foi escolhido como objeto deste estudo. O foco principal é sua aplicação n processo ensino-aprendizagem e, embora o objeto de estudo seja o *ChatGPT*, as análises realizadas durante essa pesquisa podem ser aplicadas para outras ferramentas de inteligência artificial semelhantes.

O relatório técnico do *GPT* [12] pode ajudar a elucidar melhor o funcionamento do *ChatGPT*. Conforme apontado por [12], o *GPT* funciona por meio da geração de texto preditivo, onde o usuário insere uma pergunta ou solicitação e o modelo retorna uma resposta adequada com base no contexto fornecido. Essa tecnologia utiliza técnicas avançadas de processamento de linguagem natural e aprendizado de máquina para entender e produzir textos em linguagem natural. Tais modelos têm se mostrado capazes de realizar diversas tarefas humanas, como gerar imagens e processar textos em linguagem natural. Conforme os artigos [4] e [6], tais ferramentas apresentam, frequentemente, uma interface bastante amigável, permitindo ao usuário interagir com o modelo através de *prompts*: comandos textuais utilizando uma linguagem natural.

1.4 Objetivo Geral

Analisar a bibliografia disponível, discutido em maiores detalhes no capítulo 2 , quanto às capacidades reais do *ChatGPT* nas diferentes áreas do conhecimento, como a matemática e a programação, buscando compilar os métodos que permitem obter o melhor resultado e apontar a maior quantidade de riscos que a ferramenta pode apresentar para o processo de ensino, auxiliando assim discentes e docentes a pensar criticamente sobre o uso dessas ferramentas no processo de ensino-aprendizagem.

1.5 Objetivos Específicos

- Investigar como o *ChatGPT* e o *GPT* funcionam, analisando suas capacidades e limitações;
- Investigar como o *ChatGPT* é aplicado ao processo ensino-aprendizagem;
- Identificar os benefícios proporcionados pelo uso do *ChatGPT* no processo ensino-aprendizagem;

- Apontar os desafios e riscos associados ao uso do *ChatGPT* no ensino e na pesquisa acadêmica;
- Apontar recomendações para redução dos riscos associados ao uso do *ChatGPT* no ambiente acadêmico;

1.6 Estrutura do Trabalho

O presente trabalho divide-se em 10 capítulos, abordando desde os conhecimentos mais básicos sobre inteligências artificiais até estudos de caso envolvendo o uso dessas ferramentas no ambiente educacional. Os capítulos estão organizados da seguinte forma:

- O capítulo 3 busca elucidar conceitos fundamentais relacionados ao estudo e desenvolvimento de inteligências artificiais, introduzindo o leitor aos conhecimentos necessários para entender como as inteligências artificiais desenvolvidas atualmente funcionam e o que conseguem ou não fazer;
- O capítulo 4 explora a evolução do *ChatGPT*, desde sua criação até as versões mais recentes, destacando as capacidades e limitações apresentadas pelo modelo;
- O capítulo 5 discute o uso da inteligência artificial no processo ensino-aprendizagem, incluindo a personalização do ensino, a agilidade nas respostas, e os desafios éticos que surgem com sua implementação;
- O capítulo 6 apresenta os benefícios do uso do *ChatGPT*, como a criação de avaliações de aprendizagem e a oferta de tutoria personalizada, que podem melhorar a experiência educacional;
- O capítulo 7, em contrapartida, aborda os riscos associados ao uso dessa tecnologia, incluindo riscos a integridade acadêmica e a disseminação de informações imprecisas;
- O capítulo 8 propõe recomendações para mitigar os riscos identificados, como a implementação de regras claras e a promoção de capacitações para docentes e discentes.
- O capítulo 9, são discutidos os resultados da pesquisa, no qual se analisa a eficácia do *ChatGPT* no ambiente acadêmico;
- O capítulo 10 apresenta as conclusões do trabalho, ressaltando a importância de um uso responsável e ético do *ChatGPT* no processo ensino-aprendizagem;
- O capítulo 11 apresenta trabalhos futuros relacionado ao tema do presente artigo.

Capítulo 2

Metodologia

A pesquisa foi conduzida por meio de uma revisão de literatura baseada em artigos científicos, livros e relatórios técnicos sobre o tema. A análise também incluiu a comparação de diferentes perspectivas acadêmicas sobre o uso do *ChatGPT* no processo ensino-aprendizagem.

A base de dados utilizada para obtenção da literatura citada foi o portal de periódicos da fundação CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), uma vez que o acesso aos seus periódicos é fornecido de forma gratuita para todos os estudantes da Universidade de Brasília. Foram utilizados, também, apenas artigos publicados entre os anos de 2022 (ano de lançamento do *ChatGPT*) e 2023 (ano em que foi realizada esta pesquisa).

A entrada utilizada para buscar as publicações analisadas foi "*ChatGPT for education benefits and risks*". Foram descartadas as publicações que não eram relacionados à educação como, por exemplo, jornalismo e riscos legais. Além daquelas que não traziam benefícios e riscos associados ao processo de ensino-aprendizagem. Ao final, foram escolhidos os seis primeiros artigos que abordavam o tema do uso do *ChatGPT* no processo ensino-aprendizagem sob a ótica de seus benefícios e riscos: [13], [7], [11], [14], [15] e [16].

Ao longo das pesquisas para documentação dos conceitos utilizados para fundamentar as discussões sobre as ferramentas baseadas em inteligência artificial, bem como de suas capacidades e limitações, tomou-se ciência de [17] e [18], que apresentaram um bom retrato do uso dessas ferramentas no Brasil e passaram a figurar entre os trabalhos escolhidos para compor esta revisão.

2.1 Limitações

Este trabalho não compreende um levantamento bibliográfico amplo, mas sim uma compilação de artigos que fizeram uma análise bibliográfica dos riscos e benefícios. A escolha

limitada de artigos impõe uma limitação estatística aos resultados deste trabalho, uma vez que pesquisas relevantes que não apareceram nas primeiras posições devido a fatores como algoritmos de ranqueamento, palavras-chave utilizadas ou atualizações recentes na base de dados podem ter sido excluídas.

O número reduzido de publicações escolhidas para compor esta pesquisa pode comprometer a abrangência e a representatividade dos estudos analisados. Ademais, o uso de inteligências artificiais como o *ChatGPT* é recente e seus efeitos ainda estão por ser descobertos. Por essa razão, os resultados obtidos ao final deste trabalho não representam uma listagem exaustiva dos benefícios, riscos e recomendações.

Capítulo 3

ChatGPT

O capítulo apresenta uma introdução abrangente ao funcionamento do *ChatGPT*, explorando conceitos fundamentais que sustentam a inteligência artificial e o aprendizado de máquina. Ele descreve o papel do *ChatGPT* como uma ferramenta baseada em modelos de linguagem de grande escala e destaca seu impacto em diversas áreas, como saúde, educação e matemática. Além disso, o texto aborda tópicos relacionados à confiabilidade e às limitações do *ChatGPT*, enfatizando a necessidade de compreensão dos princípios subjacentes para um uso eficaz. Ao longo do capítulo, conceitos como inteligência artificial 3.1, aprendizado de máquina 3.2, viés algorítmico 3.3, aprendizado profundo 3.4, *LLMs* 3.5 e representações de informações são detalhados, oferecendo uma base teórica para o entendimento da tecnologia por trás do *ChatGPT* e suas implicações no mundo moderno.

O *ChatGPT* é uma ferramenta disponibilizada a partir do *site* <https://chat.openai.com/>. O *site* fornece acesso a uma caixa de texto que recebe entradas em formato natural e utiliza inteligência artificial para fornecer respostas para perguntas ou comandos. É possível, inclusive, perguntar ao *ChatGPT* "o que é *ChatGPT*?", pergunta para a qual obtém-se a resposta mostrada na imagem 3.1. Muito além de um simples site, o *ChatGPT* tem se mostrado útil em diversas áreas do conhecimento humano.

Estudos como [19] têm examinado a capacidade do *ChatGPT* em oferecer recomendações de tratamento para doenças pediátricas. Embora demonstre promessa ao fornecer sugestões de tratamento gerais para condições pediátricas comuns, aconselha-se cautela ao considerar seu uso para doenças raras. Além disso, [20] destaca que o *ChatGPT* foi avaliado por seu desempenho diagnóstico em radiologia, mostrando potencial para auxiliar no diagnóstico diferencial.

Estudos dentro do setor educacional, como os de [21], demonstram que o *ChatGPT* tem sido explorado por seu papel no apoio aos processos de aprendizado de idiomas. Pesquisas indicam que ele pode auxiliar os alunos na conclusão de tarefas de escrita, destacando o seu potencial como ferramenta educacional. Além disso, [22] avalia o *ChatGPT* por

suas capacidades matemáticas, sugerindo que pode apoiar matemáticos profissionais em tarefas como resposta a perguntas e busca de teoremas.

Conforme [23], avaliações foram realizadas para analisar a confiabilidade do *ChatGPT*. Descobertas têm mostrado que, embora o *ChatGPT* possa fornecer informações precisas, também pode apresentar informações significativamente imprecisas aos usuários. Além disso, [24] destaca que avaliações críticas das respostas do *ChatGPT* ao longo do tempo têm buscado identificar suposições-chave e promover a literacia em inteligência artificial generativa.

ChatGPT: Modelo PLN Poderoso

February 28, 2024



Anonymous

O que é ChatGPT?



ChatGPT

O ChatGPT é um modelo de linguagem desenvolvido pela OpenAI, baseado na arquitetura GPT-3.5 (Generative Pre-trained Transformer 3.5). Trata-se de um modelo de aprendizado de máquina que utiliza técnicas de processamento de linguagem natural (PLN) para gerar respostas coesas e contextualmente relevantes a partir de entradas de texto.

O "GPT" em ChatGPT refere-se à abordagem "Generative Pre-trained Transformer", que destaca a capacidade do modelo de gerar texto automaticamente após ter sido treinado em uma grande quantidade de dados textuais. O número "3.5" indica a versão específica da arquitetura GPT usada.

O treinamento prévio do modelo envolve expô-lo a uma diversidade de textos em vários contextos, permitindo que ele adquira um entendimento amplo da linguagem. Isso possibilita que o ChatGPT responda a uma variedade de perguntas, forneça informações, crie texto fictício, entre outras tarefas relacionadas à linguagem natural.

Vale ressaltar que o ChatGPT é uma ferramenta poderosa, mas seu conhecimento está limitado ao que foi treinado até a minha última atualização em janeiro de 2022, e não tem conhecimento em tempo real.

Fonte: <https://chat.openai.com/share/bca30ca9-f0bd-484f-adcf-9ebb24cf3b2e>

Figura 3.1: Resposta do ChatGPT para a pergunta "o que é ChatGPT". Observe que além de responder, o ChatGPT ainda gera um título para a interação: "ChatGPT: Modelo PLN Poderoso".

É possível obter várias informações a partir dessa resposta. De fato, conforme exposto

por [6], o *ChatGPT* é baseado no *GPT*, um grande modelo de linguagem, constituindo uma ferramenta focada no processamento de linguagem natural através do aprendizado de máquina. Fazendo, portanto, parte do conjunto de ferramentas baseadas em inteligência artificial como afirma [25, 26]. Pode ser mais fácil visualizar a relação entre os conceitos citados a partir de um Diagrama de Venn 3.2.

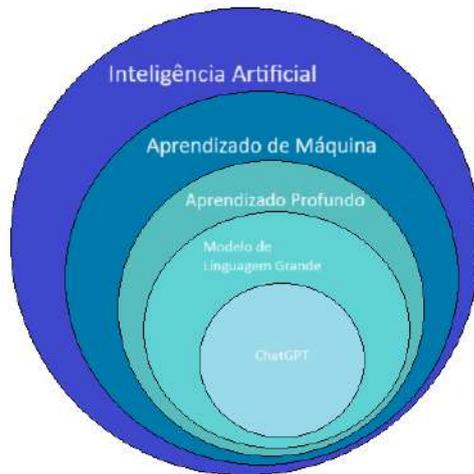


Figura 3.2: Relação entre os conceitos utilizados no campo da inteligência artificial

A resposta do *ChatGPT* deixa claro que, para que seja possível entender o seu funcionamento, é necessário esclarecer alguns conceitos fundamentais como: inteligência artificial, aprendizado de máquina, grandes modelos de linguagem, entre outros. Este capítulo tem como objetivo fazer uma introdução a cada um desses conceitos essenciais, para que seja possível entender como o *ChatGPT* processa os comandos e produz as respostas. O capítulo seguinte focará em características próprias do *ChatGPT* e seu modelo embasador: o GPT.

3.1 Inteligência Artificial

Define-se, segundo [25], que a inteligência artificial (IA) é um elemento central na ciência da computação, e, segundo [27], tem se tornado uma técnica dominante para a solução de problemas. Trata-se de um desenvolvimento científico que envolve a criação de máquinas capazes de realizar ações que se enquadram no escopo da inteligência humana,

como ensino, lógica, autocorreção e autoprogramação. Ainda segundo [25], o objetivo principal da inteligência artificial é capacitar os computadores a executarem tarefas que tradicionalmente seriam realizadas por humanos. Atualmente, ela permeia as atividades diárias das pessoas, incluindo indústria, agricultura e medicina. As habilidades que a inteligência artificial pode aprimorar incluem aprendizado, reconhecimento de padrões, raciocínio, resolução de problemas, percepção visual e compreensão de linguagem. Além disso, a pesquisa em inteligência artificial é altamente técnica, especializada e depende fortemente de lógica, conhecimento, planejamento, educação, comunicação, imagens e da capacidade de mover e manipular objetos.

A inteligência artificial configura um interesse da ciência da computação datável desde, pelo menos, 1950, com a publicação por Alan Turing de *Computing Machinery and Intelligence* [28]. Em seu artigo, Turing descreve um jogo chamado Jogo da Imitação (*Imitation Game*), cujo objetivo principal era determinar se computadores podem pensar ("*Can machines think?*"). Este jogo ficou conhecido como Teste de Turing (*Turing Test*) e, como observado por [29], embora seja muito útil como um exercício de pensamento, não teve seu contexto bem definido por seu autor.

Estudos como os realizados por [29] indicam que a atual versão do GPT-4 seria capaz de passar no entendimento popular do que seria o Teste de Turing. A própria proposta do teste se relaciona bastante bem com interpretações modernas do que seria a Inteligência Artificial. Para facilitar a exploração dos conceitos, este artigo adotará uma das definições apresentadas pelo diário *Transactions on Artificial Intelligence* [30] da *IEEE*, que define Inteligência Artificial como a automação da cognição (*Artificial Intelligence is the automation of cognition*).

3.2 Aprendizado de Máquina

Traduzido do inglês *Machine Learning* (ML), compõe, segundo [25, 26], uma subárea da inteligência artificial. De acordo com [27] e [26], tal área é especialmente focada em algoritmos capazes de "aprender" a partir da inserção de novos dados. Ao contrário do desenvolvimento de software convencional, segundo [25] e [26] a aprendizagem de máquina se dedica a criar algoritmos capazes de aprender de forma autônoma a partir de dados. De fato, o aprendizado de máquina é definido por [31] como o desenvolvimento de algoritmos capazes de fazer previsões precisas através da análise de grandes conjuntos de dados, sendo possível observar inclusive que quanto maior o conjunto de dados disponível, melhores tendem a ser os resultados.

Segundo [31], são vários os problemas tratáveis utilizando aprendizado de máquina. Dentre os problemas citados, dois são mais relevantes para o entendimento deste artigo:

classificação e regressão. Ainda segundo [31], os problemas de classificação estão relacionados à necessidade de classificar elementos de um conjunto em diferentes categorias, enquanto os problemas de regressão envolvem prever um valor real para os itens desse conjunto. Além disso, conforme observado por [5], o *ChatGPT* realiza tarefas de regressão, uma vez que, dentre as atividades executadas por ele, prever quais são as palavras mais prováveis de se seguir em uma frase está entre as mais importantes.

Um bom exercício para entender a proposta dos algoritmos de aprendizado de máquina, segundo [31], é imaginar um conjunto de documentos onde cada documento pertencente ao conjunto pode ser categorizado de acordo com um critério preestabelecido. Esse critério pode partir de várias características comuns a documentos do conjunto, como datas, autores, estilo de escrita, bem como combinações entre essas características. A partir desse conjunto, monta-se um subconjunto de documentos sobre os quais já se detém o conhecimento prévio da categoria a qual pertencem, já foram rotulados. Trata-se do "conjunto de treinamento", conjunto esse definido por [32] como um subconjunto usado nos estágios iniciais do desenvolvimento do aprendizado de máquina, responsável por ajustar parâmetros internos do algoritmo.

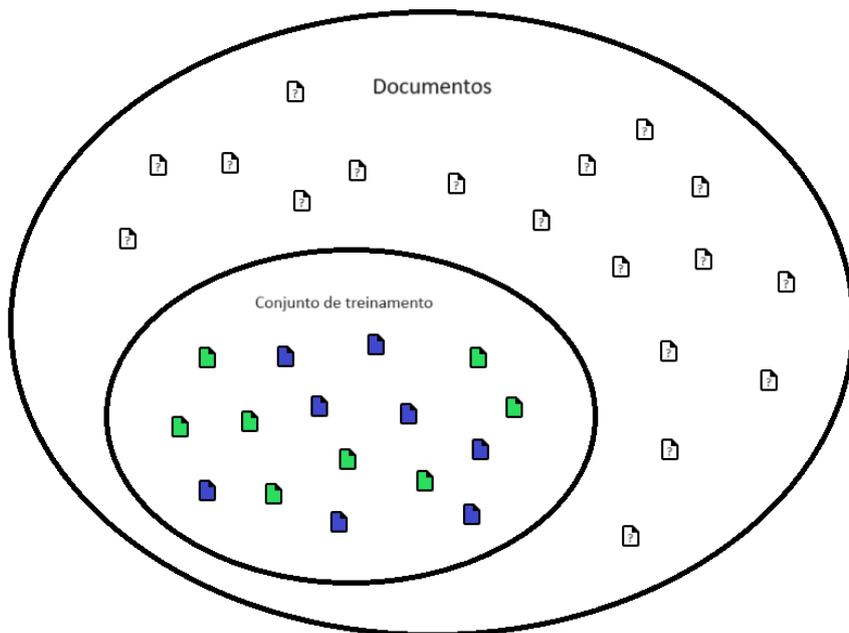


Figura 3.3: Visualização dos conjuntos de documentos baseado no exemplo apresentado em *Foundations of Machine Learning*[31].

Este exemplo constitui um tipo específico de aprendizado de máquina: o aprendizado supervisionado. Neste modelo, como explicado em [31], o conjunto de treinamento já vem rotulado, ou seja, já se determina a saída esperada para cada entrada, sendo frequentemente responsabilidade de um supervisor humano conferir as respostas, de acordo com [26]. Existem também modelos não supervisionados como os citados por [31] e [26] que

os definem como modelos nos quais o conjunto de treinamento é montado sem possuir qualquer tipo de rótulo e cabe unicamente ao algoritmo estabelecer as relações entre os elementos do conjunto. O processo de rotulagem, do inglês *labeling*, seria, no exemplo dos documentos, a atividade humana de categorizar os documentos conforme a característica escolhida e é a diferença fundamental entre os dois modelos.

Escolhido o conjunto de treinamento, a próxima etapa, segundo [31], é alimentar o algoritmo desenvolvido com ele. Este processo é, de acordo com [33], essencial para permitir que o modelo aprenda os padrões e relacionamentos subjacentes dentro dos dados, além de constituir a base para o processo de tomada de decisão do modelo. Além disso, Trata-se, conforme apontado por [34], de um momento para otimizar o modelo em relação ao consumo de recursos, como uso de memória no dispositivo de destino, após a fase de treinamento. Voltando ao exemplo dos documentos, o processo de treinamento regularia os parâmetros internos do algoritmo permitindo que ele cruzasse informações relativas aos documentos de forma a criar uma generalização de como seria um documento pertencente a cada categoria. Segundo [31], a partir dessa generalização, o algoritmo passa a ter uma probabilidade cada vez maior de prever a qual categoria pertenceria um documento desconhecido.

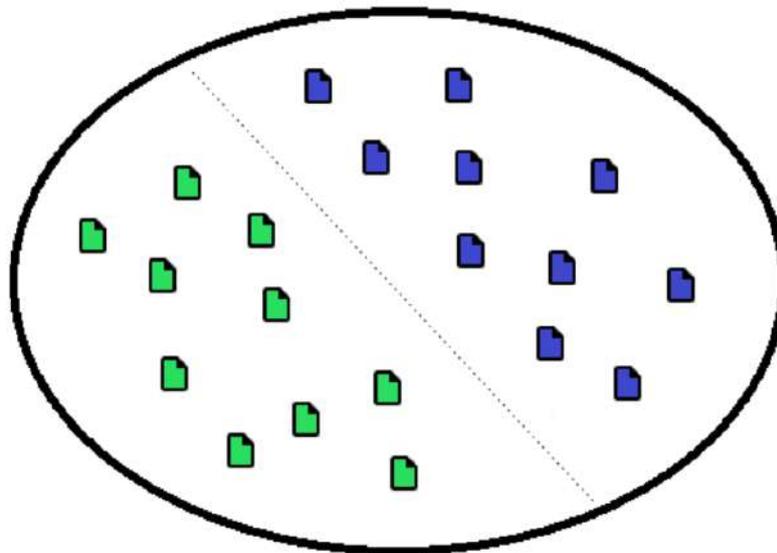


Figura 3.4: Abstração da organização dos arquivos dentro do algoritmo após treinamento. a linha pontilhada é o resultado do balanceamento dos parâmetros.

De fato, o próprio [31] afirma que o objetivo final de um algoritmo de aprendizado de máquina é criar generalizações sobre os objetos analisados. Interpretando os conceitos trazidos no livro [31] de uma forma bem simples, um algoritmo de aprendizado de máquina é composto por um conjunto de parâmetros estudados e mapeados para atributos reais dos

objetos focos da análise. A esses parâmetros se ligam pesos numéricos que representam a relevância do parâmetro na classificação do objeto. A cada rodada de treinamento, os pesos são atualizados para refletir melhor a coisa representada e, assim, melhorar a capacidade preditiva do modelo. O problema dessa dependência dos dados é que, segundo [35], se os dados de treinamento refletirem preconceitos históricos ou desigualdades sociais, os algoritmos provavelmente reproduzirão esses vieses em suas previsões e decisões, gerando um viés algorítmico 3.3.

Entretanto, de acordo com [26], algoritmos de aprendizado de máquina são, extremamente dependentes da representação da informação que eles recebem. Dado o exemplo dos documentos, para que se faça uma classificação precisa é necessário estabelecer parâmetros que descrevem bem cada um dos documentos, assim como criar um padrão de codificação que faça uma boa representação dos documentos. Um bom exemplo do que seria um padrão de codificação é o formato *BibTex* comumente usado para fazer citações ao se utilizar a linguagem *LaTex* para escrita de artigos acadêmicos.

```
@article{alan1950a,  
  title = {Computing Machinery and Intelligence},  
  year = {1950},  
  journal = {Mind},  
  pages = {433-460},  
  volume = {49},  
  author = {Turing, Alan Mathison}  
}
```

Figura 3.5: Exemplo de codificação da publicação *Computing Machinery and Intelligence* de Alan Turing no formato *BibText*.

É interessante notar que o problema da representação da informação não é uma preocupação da ciência da computação, mas pode ser observado mesmo em situações comuns à vida diária. Um exemplo desse fenômeno, citado por [26], é a facilidade em realizar operações aritméticas usando numerais arábicos, mas se tratar de uma tarefa tão exaustiva em numerais romanos que foi necessário a introdução do ábaco para facilitar a realização de operações usuais no dia a dia.

3.3 Viés Algorítmico

Os estudos dispostos em [35] trazem diversas definições interessantes para entender o conceito de viés algorítmico, bem como de suas consequências. O primeiro conceito trazido é o de viés (em inglês, *bias*), termo usado para referir-se, de modo geral, a uma inclinação ou preconceito a favor ou contra um objeto, pessoa ou grupo em relação a outro, geralmente de forma considerada injusta. Em seguida, os autores trazem o conceito de viés algorítmico: falhas sistemáticas e repetitivas em algoritmos de computador que resultam em desfechos injustos, como favorecer um grupo de usuários em relação a outro.

De acordo com [36], a capacidade de adaptação dos algoritmos de aprendizado de máquina permite sua aplicação em vários campos de atuação, como contratação, educação, análise de crédito, justiça criminal e negociação de ações. Infelizmente, os autores ressaltam que essa adaptabilidade permitiu a emergência de novos e complexos tipos de discriminação. Discriminação essa que pode surgir de várias etapas que constituem o desenvolvimento de um algoritmo de aprendizado de máquina, conforme aponta [36].

Ainda segundo [36], o viés algorítmico pode ser resultante do enviesamento dos dados utilizados para treinar o algoritmo. Isso pode ocorrer quando os dados utilizados refletem discriminações praticadas no passado. Para consolidar esse entendimento, os autores trazem o exemplo de um algoritmo de contratação que utiliza informações de antigos funcionários de uma empresa. Esse algoritmo pode aprender a repetir padrões preconceituosos que resultaram em pouca representação de grupos minoritários.

Além disso, segundo [35], o viés algorítmico pode surgir a partir do design e da implementação dos próprios algoritmos. Fatores como a seleção de características, a arquitetura do modelo e a aplicação contextual dos algoritmos podem introduzir vieses que não estão presentes nos dados brutos.

É importante, também, se atentar às observações feitas em [37], onde salienta-se o risco de consequências do "mundo real"(em inglês, *real-world*) proporcionadas pelo viés algorítmico. Dois tipos de vieses algorítmicos são definidos por [37]: viés de gênero (em inglês, *gender bias*) e viés racial (em inglês, *racial bias*). Um exemplo de viés de gênero trazido por [37] é o de uma ferramenta de recrutamento utilizada pela Amazon, a qual passou a exibir um viés que desfavorecia a contratação de mulheres.

3.4 Aprendizado Profundo

A metodologia básica de aprendizado de máquina apresentada anteriormente obteve bastante sucesso em atividades nichadas como, conforme afirma [26], aprender a jogar xadrez ou identificar se um *e-mail* trata-se de um *spam*. Mas, ainda segundo [26], mesmo ativida-

des fáceis e intuitivas para seres humanos, como reconhecer palavras faladas ou diferenciar rostos em fotografias, ainda se mostravam desafiadoras para algoritmos de aprendizado de máquina mais simples. A grande dificuldade, como afirma [26], encontra-se na representação da informação a ser fornecida para o algoritmo. Os algoritmos de visão computacional, por exemplo, cujo objetivo é reconhecer padrões como rostos ou texto em imagens, vão ter, muitas vezes, como objeto de análise, uma sequência de *bytes* que codificam uma organização de cores e posições de *pixels*.

Segundo [26], o problema é que, apesar de ser perfeitamente possível produzir algoritmos de aprendizado de máquina eficazes, essa eficácia será restringida pela capacidade de representar bem os dados a serem fornecidos para o algoritmo. Não sendo o bastante, é comum, conforme observado por [26], haver dificuldades em encontrar boas representações mesmo para problemas simples. Um algoritmo de visão computacional cujo objetivo é identificar carros em fotografias, por exemplo, poderia ser programado para identificar a presença de rodas, por exemplo, entretanto é muito difícil representar o que seria uma roda em termos de *pixels*.

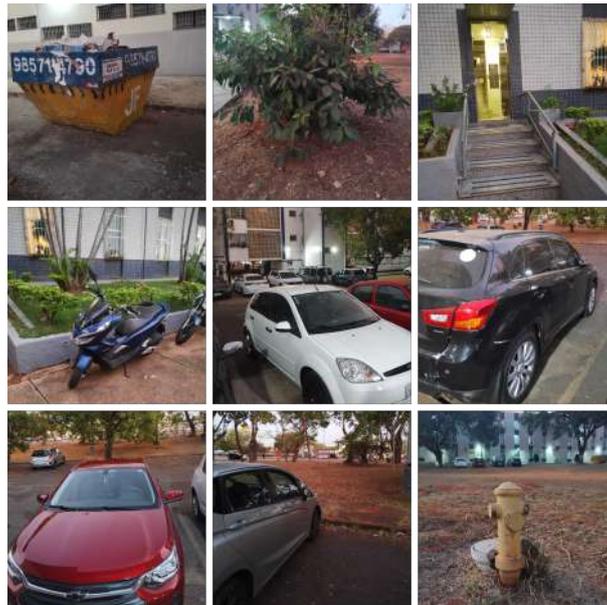


Figura 3.6: Imagem simulando um *captcha*, ferramenta usada para identificar se um usuário é humano. A ideia é justamente que é difícil para um computador conseguir compreender que imagens contém a representação de um carro.

O aprendizado profundo, destaca [26], busca resolver o problema da representação ao usar representações formadas por novas representações mais simples, tendo seu melhor exemplo em um algoritmo de chamado Perceptron Multicamadas (do inglês *Multilayer Perceptron*), ou *MLP*. Apesar do nome desafiador, constitui um modelo bastante simples, consistuindo em uma função matemática mapeando uma entrada a uma saída. O

diferencial é que essa função é formada a partir da combinação de várias outras funções matemáticas menores. Esta composição de funções para representar os diferentes níveis de informação é um dos pontos centrais do aprendizado profundo.

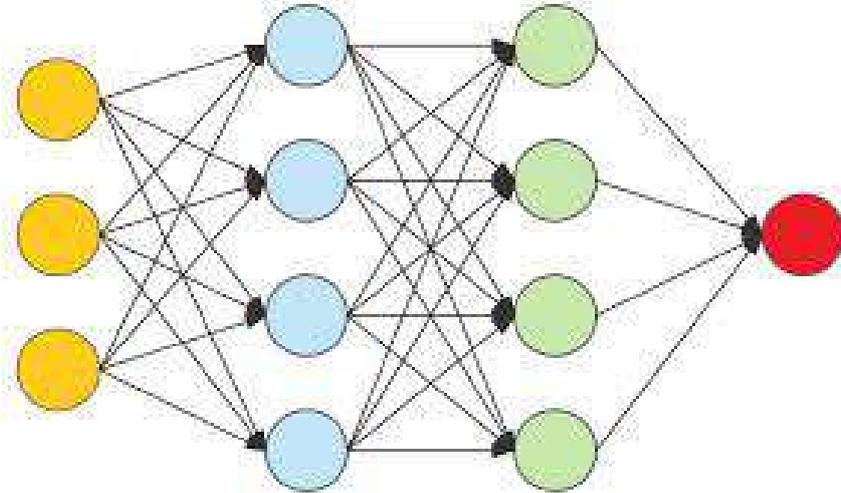
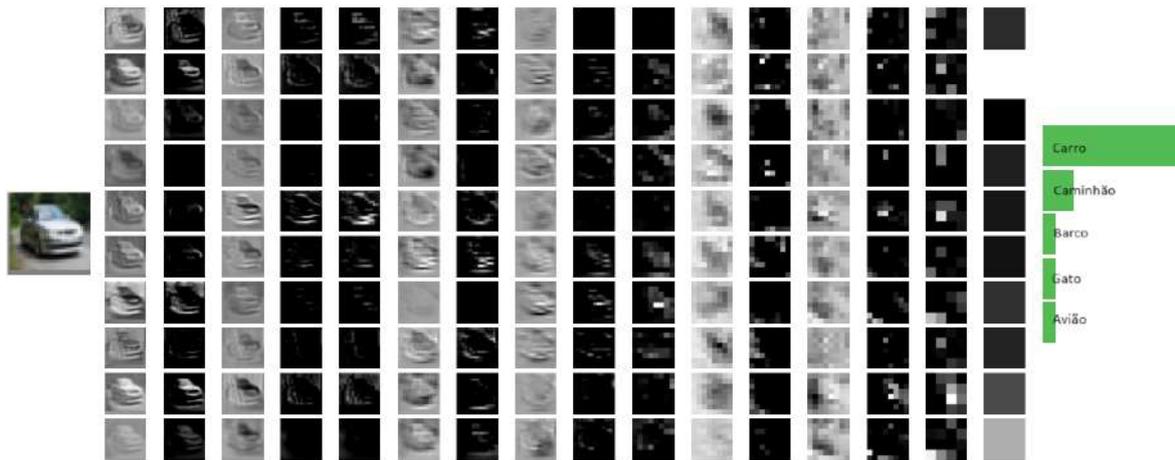


Figura 3.7: Representação de uma Perceptron Multicamadas (*MLP* em um grafo ordenado. Cada vértice do grafo representa um perceptron, enquanto as arestas representam as entradas e saídas de cada perceptron.

O aprendizado profundo é definido por [38] como um modelo computacional que envolve múltiplas camadas de processamento para aprender representações de dados com vários níveis de abstração. Segundo [39], não representa apenas uma ferramenta poderosa para análise de dados, mas também um avanço metodológico em inteligência artificial capaz de descobrir relações importantes e não intuitivas em conjuntos de dados complexos. Pode ser difícil para um computador entender informações sensoriais (como imagens, sons, cheiros) crua, neste cenário, é de entendimento de [26] que dividir a informação em porções menores permite construir conceitos complexos a partir de conceitos mais simples.

Segundo [40, 41], tal funcionamento foi pensado para o modelo de aprendizado profundo ao tentar simular o próprio cérebro humano. De fato, as semelhanças entre a forma como algoritmos de aprendizado profundo aprendem, através da experiência, sem depender unicamente da programação estática e previsão de todos os cenários possíveis, como os algoritmos convencionais, levantou interesse entre neurocientistas como afirmado em [41]. Recentemente, é possível encontrar uma vasta literatura comparando o funcionamento do cérebro humano ao de sistemas operacionais. De acordo com [41], embora não haja consenso sobre a possibilidade de compreender o funcionamento do cérebro através de algoritmos de aprendizado profundo, existem boas evidências de que uma arquitetura bem desenvolvida de tais modelos poderia aproximar o funcionamento dos algoritmos aos do cérebro humano.



Fonte: <http://cs231n.stanford.edu>

Figura 3.8: Imagem representando um modelo de aprendizado profundo onde cada coluna representa uma das camadas de representação da informação.

3.5 *LLMs*

LLM - Large Language Model (Modelo de Linguagem Grande) - refere-se, como observado por [42], ao termo cunhado pela comunidade científica para se referir a modelos de linguagem que possuem um número significativo de parâmetros (dezenas ou centenas de bilhões de parâmetros). *LLMs*, cita [43], compreendem modelos de linguagem que utilizam técnicas de aprendizado profundo associadas a uma grande quantidade de informação. Assim como os algoritmos de aprendizado de máquina citados por [31], é de entendimento de [44] que os *LLMs* demonstraram uma eficiência diretamente proporcional ao seu número de parâmetros e quantidade de informação usada em seu treinamento. Tais parâmetros, cita [5], são o resultado da distribuição da informação em um espaço vetorial, os quais, uma vez distribuídos, podem ser utilizados por algoritmos estatísticos para prever que palavras ou frases têm a maior probabilidade de aparecerem juntas. Esses modelos de linguagem foram, conforme descrito por [42, 5], utilizados para executar diversas tarefas relacionadas ao *NLP - Natural Language Processing* (Processamento de Linguagem Natural) - que busca permitir aos computadores executar tarefas descritas a partir de uma linguagem humana.

Segundo [45], desde o seu surgimento, um dos principais objetivos da inteligência artificial tem sido avançar nos métodos computacionais para a compreensão da linguagem natural. Conforme [46], o *NLP* compreende de um ramo da Inteligência Artificial cujo início remonta a 1950 e, de acordo com [47], refere-se a computadores e sistemas que analisam, programam ou produzem em uma ou mais línguas humanas. Compreende

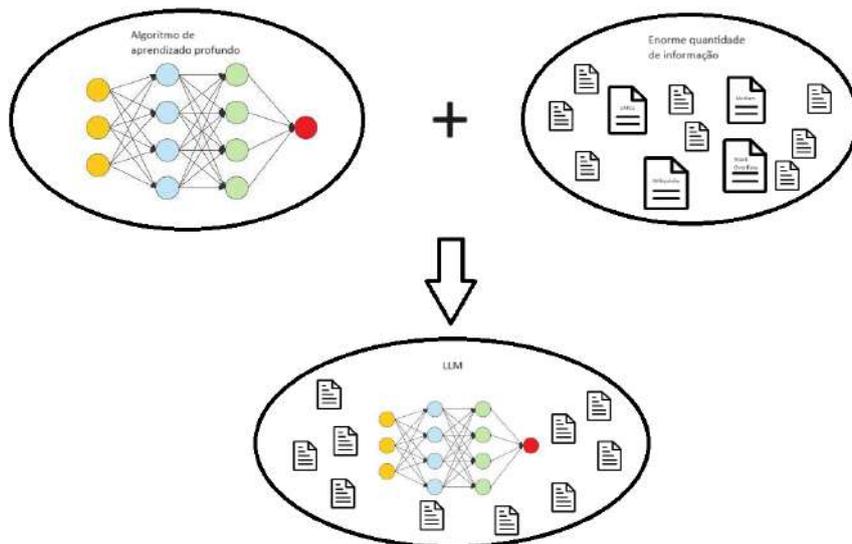


Figura 3.9: *LLMs* como o resultado do treinamento de algoritmos de aprendizado profundo utilizando grandes quantidades de dados.

ainda, como destacam [47] e [48], um campo de ampla pesquisa na área de inteligência artificial e da linguística que se concentra na interação entre computadores e linguagem humana, com objetivo de capacitar as máquinas a entender, interpretar e gerar texto ou fala de maneira semelhante à comunicação humana, que por sua vez possui vários tipos de compreensão e níveis de combinações linguísticas.

De acordo com [44], o processamento de linguagem natural teve seu início marcado pelo uso de redes neurais, sendo posteriormente substituído por modelos de linguagem pré-treinados (em inglês, *pre-trained language models* - *PLM*. Modelos de linguagem (em inglês *Language Model* - *LM*, já haviam sido utilizados com sucesso para atividades de cunho específico, mas diferem dos modelos pré-treinados por dependerem de um modelo de aprendizado supervisionado, em oposição ao modelo auto-supervisionado utilizado pelas *PLMs*. O aumento na quantidade de informações e parâmetros utilizados pelas *PLMs* foi o responsável pela transição para modelos grandes, os *LLMs*.

Segundo [49], o modelo auto-supervisionado se diferencia bastante dos, já mencionados, modelos supervisionado e não supervisionado, principalmente em relação ao tratamento dos dados que serão usados no treinamento. Foi considerado por [49] uma boa solução para um problema que ameaçava estagnar o algoritmo de aprendizado profundo: a rotulagem. Segundo [44], embora a quantidade de informação produzida atualmente tenha dado o insumo necessário para algoritmos bastante eficientes, os algoritmos que dependiam de um modelo de aprendizado supervisionado, como os *LLMs*, exigiam, de acordo com [49], um grande tempo de trabalho humano para fazer as rotulagens, uma vez que a quantidade de informação a ser categorizada era enorme. A solução, afirma [49],

foi utilizar um conjunto de treinamento cuja rotulagem havia sido feita através de um algoritmo de aprendizado profundo não supervisionado.

Tipicamente, os *LLMs* se referem a modelos de linguagem transformadores, como afirmado em [42]. De fato, GPT, foco de estudo desse artigo, trata-se de uma sigla para *Generative Pre-Trained Transformer* (Transformador Generativo Pré-Treinado) e é composto por uma série de modelos de linguagem desenvolvidos pela OpenAI que demonstram ser um avanço significativo às outras abordagens anteriores do NPL como Recurrent Neural Networks (RNN) e Convolutional Neural Networks (CNN). Conforme observado por [50], ele utiliza um mecanismo de autoatenção para permitir que o modelo considere o contexto de toda a frase ao gerar a próxima palavra, o que melhorou significativamente a geração e interpretação da linguagem.

Capítulo 4

GPT

No presente capítulo, é explorada a evolução e o impacto dos modelos de linguagem da *OpenAI*, no ambiente acadêmico. A partir da sua criação em 2015 até o desenvolvimento do *GPT-4*, em 2023, será discutido como esses modelos foram treinados e evoluíram em termos de parâmetros e capacidades. Será abordada a breve história 4.1, como é feito o pré-treinamento 4.2, a arquitetura *Transformer* 4.3, como é feita a tokenização 4.4, as capacidades e limitações 4.5 e como ele permitiu avanços significativos no processamento de linguagem natural.

4.1 História

Segundo [6], a OpenAI, empresa responsável pelo desenvolvimento da série de modelos de linguagem GPT (*Generative Pre-trained Transformer*), foi fundada em dezembro de 2015 por Sam Altman, Greg Brockman, Elon Musk, Ilya Sutskever, Wojciech Zaremba e John Schulman. A primeira versão do GPT, batizada de GPT-1, foi lançada em junho de 2018 e contava com 117 milhões de parâmetros. Em novembro de 2019, anunciou-se o GPT-2, a segunda iteração do modelo transformador, que já contava com 1,5 bilhões de parâmetros. A terceira iteração do modelo veio com o lançamento do GPT-3 em junho de 2020, o qual já contava com impressionantes 175 bilhões de parâmetros. Foi graças ao GPT-3 que, em novembro de 2022, era anunciado o *ChatGPT*, uma implementação do modelo GPT-3 treinado com dados públicos e autorizados disponíveis na internet até o momento. Atualmente, em sua versão paga, o *ChatGPT* já conta com a quarta iteração do GPT, o GPT-4, lançado em março de 2023.

4.2 Pré-treinado (*Pre-trained*)

De acordo com [51], o "*pre-trained*" no nome refere-se ao processo de treinamento inicial em grandes quantidades de dados não rotulados. Durante essa fase, o *GPT* aprende representações de palavras e frases em contextos diversos. Segundo [5], o modelo tenta prever a próxima palavra em uma sequência, dada uma janela de contexto, usando a tokenização, ou seja por meio de tokens que são as unidades mais irredutíveis possíveis. Isso gera um modelo que entende a estrutura gramatical e semântica da linguagem.

Durante o pré-treinamento, o modelo também aprende a lidar com "máscaras de atenção" para prever palavras ocultas em uma frase, como apresentado em [52]. Isso aprimora a capacidade do modelo de entender o contexto e inferir informações mesmo quando parte dos dados é oculta.

4.3 *Transformer*

O *GPT* utiliza a arquitetura *Transformer*, que, de acordo com [53] foi introduzida em 2017. Por conta de sua natureza comercial, a maior parte dos detalhes sobre a implementação do *GPT* não estão disponíveis para o público, por isso, este artigo usará o modelo *Transformer* para exemplificar o funcionamento do *GPT*. Essa arquitetura, diferente das redes neurais recorrentes, é baseada em mecanismos de autoatenção. Esses mecanismos relacionam diferentes posições de uma única sequência para calcular uma representação da sequência, permitindo que o modelo capture padrões de longo alcance em sequências de dados. Como apresentado por [53], o *Transformer* é o primeiro modelo que depende inteiramente da autoatenção para calcular representações e sua entrada e saída sem usar *RNNs* alinhados por sequência ou convolução.

O objetivo inicial de [53] era, inicialmente, melhorar os processos para a solução de problemas de tradução (do inglês *transduction*), principalmente aqueles relacionados à tradução de máquina (*machine translation*). Tais problemas, envolvendo o uso de máquinas para tradução de linguagens humanas (traduções português-inglês etc.) têm sido relacionados ao processamento de linguagem natural desde 1950, como apontado por [46], e foram foco dos testes realizados em [53].

Segundo [53], este modelo foi baseado na interação entre dois componentes principais: o codificador (*encoder*) e o decodificador (*decoder*). Entretanto, diferente de modelos comumente utilizados, que exigiam complexas redes neurais recorrentes ou convolucionais, o modelo apresentado por [53] utiliza apenas mecanismos de atenção para realizar essas interações. Estes conceitos serão explorados mais detalhadamente a seguir.

4.3.1 Codificador (*Encoder*)

De acordo com [53], o codificador no modelo *Transformer* é composto por uma pilha de camadas idênticas. Cada camada possui duas subcamadas: um mecanismo de autoatenção multi-cabeça e uma rede de alimentação totalmente conectada com *embeddings* *Position-wise*. Para facilitar essas conexões residuais, todas as subcamadas no modelo, assim como as camadas de incorporação, produzem saídas com uma dimensão de 512 tokens.

Ainda conforme [53], é explicado que no mecanismo de autoatenção multi-cabeça, cada posição no codificador pode atender a todas as posições na camada anterior do codificador. Isso permite que o modelo capture dependências entre diferentes posições na sequência de entrada. O mecanismo de autoatenção calcula uma representação da sequência relacionando diferentes posições dentro dela, sendo utilizado com sucesso em tarefas como compreensão de leitura e sumarização abstrativa.

Segundo [53], a rede de alimentação *Position-wise* é aplicada a cada posição separadamente e de maneira idêntica. Ela consiste em duas transformações lineares com uma ativação ReLU no meio. Essa rede ajuda a processar ainda mais as representações em cada posição, de acordo com.

De acordo com [53], para facilitar o fluxo de informação e melhorar o treinamento, o codificador usa conexões residuais e normalização de camada. As conexões residuais permitem que a saída de cada subcamada seja adicionada à entrada, enquanto a normalização de camada ajuda a normalizar as saídas de cada subcamada.

Conforme apontado por [53], a pilha de codificadores é composta por várias camadas, geralmente 6 no modelo base. Cada camada na pilha de codificadores executa as mesmas operações, permitindo que o modelo capture dependências e representações cada vez mais complexas à medida que a sequência de entrada é processada pela pilha.

4.3.2 Decodificador (*Decoder*)

O decodificador, de acordo com [53], no modelo *Transformer*, é composto por uma pilha de camadas idênticas, semelhante ao codificador. Cada camada no decodificador possui três subcamadas: um mecanismo de autoatenção multi-cabeça, um mecanismo de atenção multi-cabeça sobre a saída da pilha do codificador e uma rede de alimentação totalmente conectada *Position-wise*.

O mecanismo de autoatenção multi-cabeça, ainda de acordo com [53], no decodificador permite que cada posição no decodificador atenda a todas as posições na camada anterior do decodificador. Isso ajuda o decodificador a capturar dependências entre diferentes posições dentro da pilha do decodificador. O mecanismo de autoatenção calcula uma

representação da sequência do decodificador relacionando diferentes posições dentro da sequência.

Além do mecanismo de autoatenção, o decodificador também inclui um mecanismo de atenção multi-cabeça sobre a saída da pilha do codificador, como aponta [53]. Isso permite que o decodificador atenda à saída do codificador e incorpore informações da sequência de entrada durante o processo de decodificação, imitando os mecanismos de atenção típicos entre codificador e decodificador usados em modelos de sequência para sequência.

Semelhante ao codificador, o decodificador usa conexões residuais e normalização de camada para facilitar o fluxo de informações e melhorar o treinamento, conforme apontado por [53]. As conexões residuais permitem que a saída de cada subcamada seja adicionada à entrada, enquanto a normalização de camada ajuda a normalizar as saídas de cada subcamada.

Para preservar a propriedade auto-regressiva do decodificador, segundo [53], a subcamada de autoatenção na pilha do decodificador é modificada para evitar que posições atendam a posições subsequentes. Isso é alcançado mascarando (definindo como $-\infty$) os valores na entrada da função softmax que correspondem a conexões ilegais.

Segundo [53], a pilha de decodificadores no modelo *Transformer* realiza as mesmas operações da pilha de codificadores, mas também incorpora atenção sobre a saída do codificador. Isso permite que o decodificador gere uma sequência de saída elemento por elemento, consumindo os símbolos previamente gerados como entrada adicional ao gerar o próximo símbolo.

4.3.3 *Embeddings*

Embeddings são fundamentais para a arquitetura *Transformer*, como apresenta [53]. Elas são utilizadas para converter tokens de entrada e saída em vetores de uma dimensão específica. No modelo *Transformer*, *embeddings* aprendidas são empregadas para converter tokens de entrada e saída nestes vetores. Além disso, o modelo *Transformer* utiliza codificações posicionais para inserir informações sobre a posição relativa ou absoluta dos *tokens* na sequência. Isso é crucial, pois o modelo não depende de recorrência ou convolução e precisa utilizar a ordem da sequência. As codificações posicionais são adicionadas às incorporações de entrada nas bases das pilhas do codificador e do decodificador. No modelo *Transformer*, as codificações posicionais são implementadas usando funções seno e cosseno de diferentes frequências, permitindo que o modelo aprenda facilmente a atender a posições relativas. Essas codificações posicionais têm a mesma dimensão das *embeddings* e são somadas para fornecer as informações posicionais necessárias para o modelo.

4.3.4 Mecanismos de Atenção (*Attention Mechanisms*)

O mecanismo de atenção no contexto do modelo *Transformer* é um componente fundamental que permite ao modelo capturar dependências entre diferentes posições em uma sequência, de acordo com [53]. A função de atenção mapeia uma consulta (*query*) e um conjunto de pares chave-valor (*key-value*) para uma saída onde a consulta, chaves, valores e saída são todos vetores. A saída é calculada como uma soma ponderada dos valores, sendo que o peso atribuído a cada valor é calculado por uma função de compatibilidade da consulta com a chave correspondente.

Além disso, o modelo *Transformer* utiliza atenção multi-cabeça (*multi-head*), o que permite que o modelo atenda conjuntamente a informações de diferentes subespaços de representação em posições diferentes, como é apresentado em [53]. Isso é alcançado projetando linearmente as consultas, chaves e valores várias vezes com diferentes projeções lineares aprendidas e, em seguida, realizando a função de atenção em paralelo nessas versões projetadas. As saídas são então concatenadas e projetadas novamente para obter os valores finais.

O modelo *Transformer* utiliza o conceito de autoatenção (*self-attention*), que relaciona diferentes posições de uma única sequência para calcular uma representação da sequência. Esse mecanismo de autoatenção tem sido utilizado com sucesso em várias tarefas de processamento de linguagem natural, incluindo compreensão de leitura, sumarização abstrativa e modelagem de linguagem.

O mecanismo de autoatenção no modelo *Transformer* permite que cada posição na sequência atenda a todas as outras posições, possibilitando ao modelo capturar dependências de longo alcance e relacionamentos entre diferentes elementos na sequência. Isso é alcançado calculando os produtos escalares das consultas e chaves, dividindo cada um pela raiz quadrada da dimensionalidade e aplicando uma função softmax para obter os pesos nos valores. A soma ponderada resultante fornece a saída do mecanismo de atenção.

4.4 Tokenização

A tokenização é uma das primeiras etapas de transformação durante o processamento de linguagem natural, de acordo com [54], de fato, é um dos primeiros passos para se tratar os dados a serem introduzidos em modelos *Transformer*, como é apresentado por [53]. Isto se faz necessário porque, para o computador, uma entrada de texto é apenas uma sequência indistinta de caracteres, assim aponta [54]. A tokenização significa, portanto, a quebra de uma sequência de caracteres em palavras, frases, símbolos ou outros elementos significativos nos chamados *tokens*, de acordo com [55]. Como o significado de cada *token* varia conforme o algoritmo utilizado, bem como a aplicação para qual o algoritmo

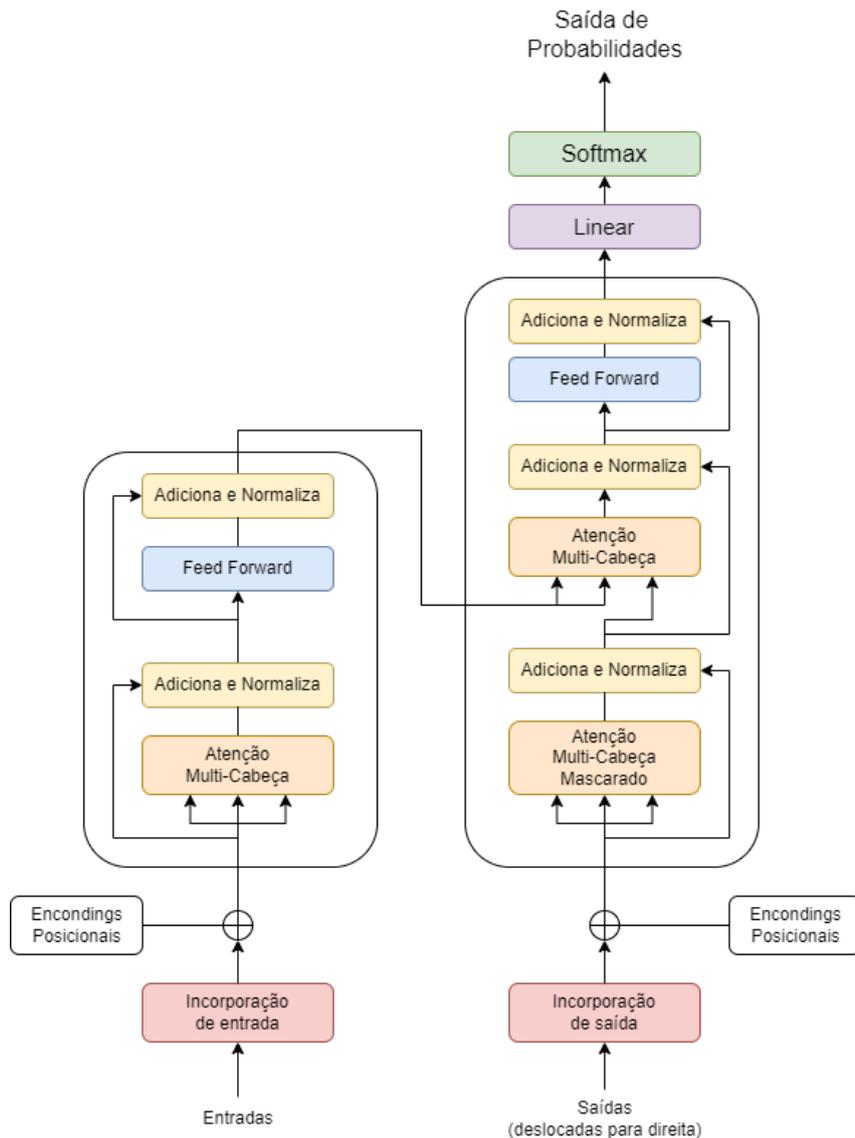


Figura 4.1: Modelo *Transformer* segundo Vaswani[53]

servirá, diferentes métodos de tokenização podem ser utilizados. Por exemplo, ao utilizar o tokenizador disponibilizado pela *OpenAI* [56] com o seguinte texto:

Teste de tokenização.

Obtemos o resultado:

"Test", "e", " de", " token", " ização"

É possível observar que o processo usado pela *OpenAI* utiliza a raiz das palavras teste ("*test*") e tokenização ("*token*"), bem como omite o ponto (".") ao final da frase. De fato, a utilização dos radicais das palavras, assim como a remoção de caracteres que podem ser considerados "ruídos" são práticas comuns no pré-processamento de texto destinado ao processamento de linguagem natural como aponta [55].

4.5 Capacidades e limitações

Em um relatório técnico publicado pela própria OpenAI [12], os desenvolvedores expuseram o GPT-4 a diversos exames originalmente pensados para seres humanos, sem qualquer treinamento específico para os testes. Embora tenha obtido resultados bastante impressionantes, o *GPT-4* demonstrou limitações similares às de seus antecessores, como aponta [12]. Ele demonstrou não ser 100% confiável, podendo sofrer as chamadas "alucinações" (erros de raciocínio e uso de fatos inventados). Os próprios autores do relatório recomendam cuidado ao usar essa ferramenta para tarefas em um contexto de alto risco, incluindo a recomendação do uso de protocolos de segurança como a revisão dos resultados por seres humanos, assim descreve [12]. Abaixo há uma tabela listando todos os exames realizados para obtenção do relatório:

Alguns resultados chamam bastante atenção e são de especial interesse para o desenvolvimento deste artigo, são eles: "*SAT Math*", "*Leetcode (easy)*", "*Leetcode (medium)*", "*Leetcode (hard)*". A prova "*SAT Math*" pode ser comparada à prova de matemática do ENEM e um resultado classificado entre os 89% melhores demonstra uma alta capacidade por parte do *GPT-4* para resolver problemas matemáticos de nível médio. *Leetcode* se refere a um site que apresenta desafios voltados para programadores que desejam testar suas habilidades ao resolver problemas algorítmicos, observa-se dos resultados a capacidade de resolver bem problemas de nível "fácil", tendo sua eficácia reduzida para problemas de nível "médio" e "difícil".

O resultado obtido no *SAT Math* pode ser ratificado por [22], o qual procura testar as capacidades do *ChatGPT* para questões de nível médio e superior. Para este estudo, os autores criaram e apresentaram um novo *dataset* para testagem de inteligências artificiais: GHOST. Este *dataset* contém uma coleção de questões de nível médio e superior, obtidas de livros usados comumente em graduações, questões formuladas especificamente para o teste por especialistas e questões obtidas do *dataset MATH*, dentre outras apresentado por [22]. Os autores ranquearam seus resultados de 1 a 5, onde 5 representa uma resposta perfeitamente correspondente à pergunta, 1 representa uma resposta completamente inadequada e 2 demonstra que a pergunta foi pelo menos compreendida. O resultado de 3,5, média entre o pior e melhor resultado, foi utilizado como limite entre sucesso e falha.

Os resultados obtidos por [22], ao aplicar o *GHOST* e *miniGHOST* (datasets menores baseados nos GHOSTs) sobre o *ChatGPT*, demonstram um resultado médio de 3,20 para o *GPT-3*, 3,29 para o *GPT-3.5* e 4,15 para o *GPT-4*. Com exceção do *GPT-3*, todas as versões testadas obtiveram *scores* acima de 3,5 para os testes do *dataset MATH*, como aponta [22], demonstrando capacidades suficientes para responder questões de nível médio e de graduação (álgebra, contagem e probabilidade, pre-álgebra e pre-cálculo). Os modelos GPT-3 e GPT-3.5, entretanto, obtiveram *scores* majoritariamente abaixo de

3,5 para questões obtidas de livros didáticos de nível superior e integrações simbólicas. Todos os modelos parecem performar bem (*scores* majoritariamente acima de 3,5) para questões envolvendo obtenção de definição e correção o compleição de provas matemáticas incorretas ou incompletas (*Hole-in-Proofs*) [22]. Os autores determinam, portanto, que o *ChatGPT* não se encontra pronto para entregar provas matemáticas de alta qualidade, sendo o *GPT-4*, contudo, considerado capaz de resolver problemas matemáticos em nível de graduação, falhando em problemas de pós-graduação. Os resultados foram sumarizados pelos autores em um gráfico, obtido a partir da publicação:

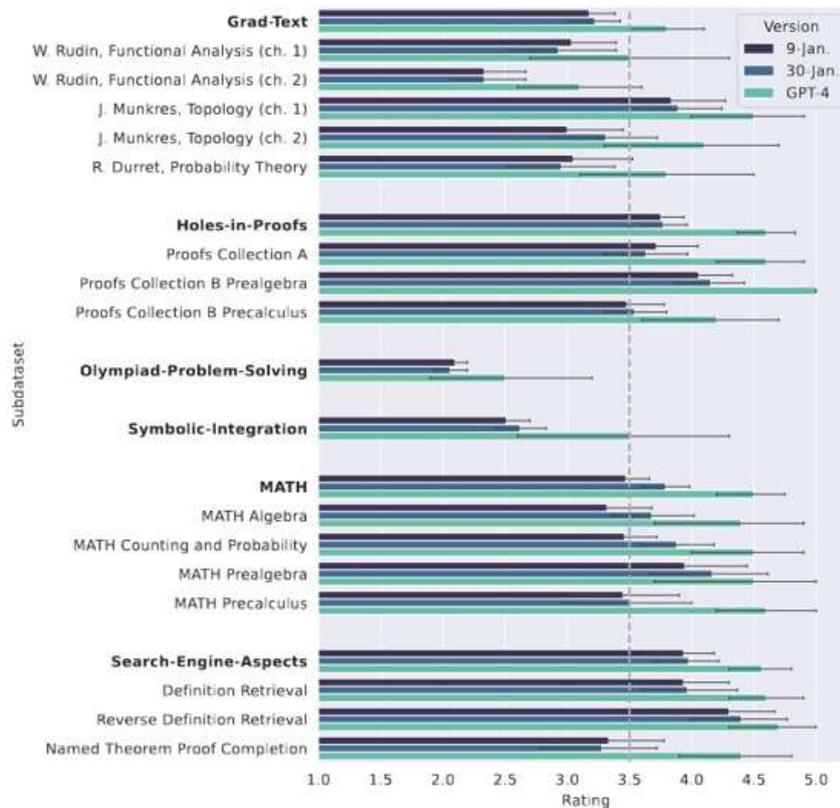


Figura 4.2: Gráfico obtido de [22]. Tradução livre da legenda original: "Classificação média para cada arquivo em cada subdataset de GHOSTS em 9 de janeiro de 2023 e as versões de 30 de janeiro de 2023 do *ChatGPT* e para miniGHOSTS no GPT-4. Observe que o máximo a classificação é 5 e a classificação mínima, onde a questão foi pelo menos compreendida, é 2, ver Apêndice B.4; a classificação mais baixa de 1 indica que a resposta erra completamente a pergunta. Assim, uma passagem razoável a nota, ou seja, 50% dos pontos, corresponde à nota 3,5, indicada pela linha pontilhada vertical. O barras de erro representam intervalos de confiança de 95%." [22]

Embora possa apresentar uma performance pior do que a de seres humanos para muitos tipos de tarefas, o *GPT-4*, segundo [12], se mostrou tão ou mais eficiente do que seres humanos para vários grupos de atividades. Entre suas conquistas, uma das mais

relevantes foi passar em um *"bar exam"*, prova prestada por aspirantes a advogados nos EUA, semelhante à prova da OAB prestada no Brasil. É importante notar que o resultado obtido pelo *GPT-4* foi classificado entre os 10% melhores, uma melhora significativa em relação ao seu predecessor, o *GPT-3.5*, que, apesar de passar no exame, ficou classificado entre os 10% piores, como demonstrado por [12].

Ao observar a tabela pode-se notar que os resultados relacionados à programação estão entre os piores obtidos pelo *GPT*. A análise feita por [57] busca revelar qualidade do código gerado pelo *ChatGPT* em sua implementação usando o *GPT-3*, como mostra [57]. Para responder à pergunta proposta no título, os autores analisam códigos gerados pelo *ChatGPT* a partir de três aspectos: corretude, compreensibilidade e segurança. Adotou-se, também, uma estratégia *"multi-round"* (multi-rodada), utilizando a habilidade do *ChatGPT* de se engajar em diálogos, como aponta [57], através da sua habilidade de reter o contexto de uma conversa e usando entradas anteriores como parte da solução do problema atual.

Como o *ChatGPT* usado para os testes usa informações produzidas até 2021 e o artigo foi publicado em agosto de 2023, os autores dividiram os problemas reunidos entre problemas anteriores a 01/01/2021 e problemas posteriores a 01/01/2021, como aponta [57]. É possível supor que problemas anteriores a 2021 estivessem presentes, mesmo que parcialmente, usados para o treinamento do *ChatGPT* até aquele momento, tornando-os mais "fáceis" para o *ChatGPT*, de acordo com [57]. De fato, os resultados mostram que o número de soluções para problemas anteriores a 2021 classificados como aceitos (68,14%) pelo juiz digital (*LeetCode*), testadas em cinco linguagens diferentes (*C*, *C++*, *Java*, *JavaScript* e *Python*), é 48,14% maior do que o número de soluções para problemas posteriores a 2021. Os resultados são bastante expressivos, visto que, mesmo para problemas de nível "fácil", o *ChatGPT* só obteve aceitação em metade (58%) das soluções apresentadas, como aponta [57].

Exames	GPT-4	GPT-4 (no vision)	GPT-3.5
Uniform Bar Exam (MBE+MEE+MPT)	298 / 400 (~90%)	298 / 400 (~90%)	213 / 400 (~10%)
LSAT	163 (~88%)	161 (~83%)	149 (~40%)
SAT Evidence-Based Reading Writing	710 / 800 (~93%)	710 / 800 (~93%)	670 / 800 (~87%)
SAT Math	700 / 800 (~89%)	690 / 800 (~89%)	670 / 800 (~70%)
Graduate Record Examination (GRE) Quantitative	163 / 170 (~80%)	157 / 170 (~62%)	147 / 170 (~25%)
Graduate Record Examination (GRE) Verbal	169 / 170 (~99%)	165 / 170 (~96%)	154 / 170 (~63%)
Graduate Record Examination (GRE) Writing	4 / 6 (~54%)	4 / 6 (~54%)	4 / 6 (~54%)
USABO Semifinal Exam 2020	87 / 150 (99% - 100%)	87 / 150 (99% - 100%)	43 / 150 (31% - 33%)
USNCO Local Section Exam 2022	36 / 60	38 / 60	24 / 60
Medical Knowledge Self-Assessment Program	75% (acertos)	75% (acertos)	53% (acertos)
Codeforces Rating	392 (< 5%)	392 (< 5%)	260 (< 5%)
AP Art History	5 (86% - 100%)	5 (86% - 100%)	5 (86% - 100%)
AP Biology	5 (85% - 100%)	5 (85% - 100%)	4 (62% - 85%)
AP Calculus BC	4 (43% - 59%)	4 (43% - 59%)	1 (0% - 7%)
AP Chemistry	4 (71% - 88%)	4 (71% - 88%)	2 (22% - 46%)
AP English Language and Composition	2 (14% - 44%)	2 (14% - 44%)	2 (14% - 44%)
AP English Literature and Composition	2 (8% - 22%)	2 (8% - 22%)	2 (8% - 22%)
AP Environmental Science	5 (91% - 100%)	5 (91% - 100%)	5 (91st - 100%)
AP Macroeconomics	5 (84% - 100%)	5 (84% - 100%)	2 (33% - 48%)
AP Microeconomics	5 (82% - 100%)	4 (60% - 82%)	4 (60% - 82%)
AP Physics 2	4 (66% - 84%)	4 (66% - 84%)	3 (30% - 66%)
AP Psychology	5 (83% - 100%)	5 (83% - 100%)	5 (83% - 100%)
AP Statistics	5 (85% - 100%)	5 (85% - 100%)	3 (40% - 63%)
AP US Government	5 (88% - 100%)	5 (88% - 100%)	4 (77% - 88%)
AP US History	5 (89% - 100%)	4 (74% - 89%)	4 (74% - 89%)
AP World History	4 (65% - 87%)	4 (65% - 87%)	4 (65% - 87%)
AMC 10 ³	30 / 150 (6% - 12%)	36 / 150 (10% - 19%)	36 / 150 (10% - 19%)
AMC 12 ³	60 / 150 (45% - 66%)	48 / 150 (19% - 40%)	30 / 150 (4% - 8%)
Introductory Sommelier (theory knowledge)	92% (acertos)	92% (acertos)	80% (acertos)
Certified Sommelier (theory knowledge)	86% (acertos)	86% (acertos)	58% (acertos)
Advanced Sommelier (theory knowledge)	77% (acertos)	77% (acertos)	46% (acertos)
Leetcode (easy)	31 / 41	31 / 41	12 / 41
Leetcode (medium)	21 / 80	21 / 80	8 / 80
Leetcode (hard)	3 / 45	3 / 45	0 / 45

Tabela 4.1: Tabela retirada do relatório publicado pela OpenAI [12]. Os resultados expressos em porcentagem representam a porcentagem de participantes abaixo do GPT, ou seja, um resultado de 90% indica que o GPT se saiu melhor do que 90% dos participantes.

Capítulo 5

Os *LLMs* (*ChatGPT*) no processo ensino-aprendizagem

Neste capítulo, é exposto como ferramentas baseadas em inteligência artificial, com destaque para o *ChatGPT*, estão transformando o ambiente acadêmico. Serão analisados seus impactos na interação entre estudantes e professores, suas aplicações práticas, como a personalização do ensino e a geração de avaliações, além de desafios éticos e técnicos, como a confiabilidade das respostas e os riscos associados à privacidade e à substituição de professores. O objetivo é oferecer uma visão abrangente sobre as possibilidades e limitações dessa tecnologia no contexto educacional, abordando também o futuro da aprendizagem com o apoio da IA.

As ferramentas baseadas em inteligência artificial estão cada vez mais presentes no ambiente acadêmico, influenciando a forma como estudantes e professores interagem e aprendem. Este capítulo explora o impacto do *ChatGPT* no ambiente acadêmico, um modelo de linguagem que utiliza aprendizado profundo para gerar respostas a partir de entradas textuais, como apresentado na sessão 3. Serão discutidos os usos dessa ferramenta, como a personalização do ensino e a agilidade nas respostas, além dos desafios éticos e práticos, como o risco de gerar dados imprecisos ou a substituição de professores. Também será analisada a aplicação do *ChatGPT* na criação de avaliações, personalização de materiais didáticos e suas implicações para o futuro do ambiente acadêmico.

Essa tecnologia tem sido aplicada em sistemas de chatbots educacionais, permitindo aos estudantes obterem respostas instantâneas para suas dúvidas e auxiliando os professores na elaboração de materiais didáticos, conforme descreve [8].

O uso de ferramentas baseadas em IA, como o *ChatGPT*, tem se tornado cada vez mais comum. Essas ferramentas, como apresentado em [10], são capazes de gerar respostas coerentes e contextualizadas a partir de um treinamento prévio com grandes quantidades de dados. O *ChatGPT* utiliza uma abordagem de aprendizado de máquina conhecida como

modelo de linguagem autoregressivo, que permite a geração de texto fluente e coerente. Ele é alimentado com exemplos de pares pergunta-resposta, onde a resposta é gerada pelo modelo com base na pergunta fornecida.

Uma das principais características do *ChatGPT* é sua capacidade de gerar respostas contextualizadas e personalizadas. Isso ocorre porque o modelo é treinado em uma ampla variedade de dados textuais, permitindo que ele aprenda padrões e contextos específicos. Além disso, o *ChatGPT* possui uma arquitetura neural avançada que permite a geração de respostas mais longas e detalhadas, como apresentado em [4]. Essa capacidade torna-o uma ferramenta útil no ambiente acadêmico, pois os estudantes podem obter dados e respostas para suas dúvidas com agilidade.

O uso do *ChatGPT* no ambiente acadêmico traz benefícios significativos, como apresenta [1]. A agilidade na obtenção de dados e respostas para dúvidas dos estudantes é um desses benefícios. Os estudantes podem interagir com o *ChatGPT* a qualquer momento e obter respostas imediatas para suas perguntas. Além disso, o *ChatGPT* permite a personalização das interações de acordo com as necessidades individuais dos estudantes. Isso significa que o sistema pode adaptar-se às preferências e estilos de aprendizagem de cada aluno, proporcionando uma experiência mais personalizada.

Segundo [58], no entanto, o uso do *ChatGPT* no ambiente acadêmico também apresenta desafios e limitações. Uma das principais preocupações é a possibilidade de o sistema gerar respostas incorretas ou imprecisas. Embora o *ChatGPT* seja treinado em grandes quantidades de dados, ele não possui capacidade para interpretar emoções ou contextos específicos. Além disso, o modelo depende de um treinamento adequado para evitar vieses ou preconceitos na geração de respostas.

As implicações éticas do uso do *ChatGPT* no meio acadêmico são preocupantes. A coleta de dados pessoais pelos sistemas de IA, necessária para fornecer respostas sob medida, levanta sérias questões sobre a privacidade dos usuários. Conforme [59] aponta que o potencial de substituição dos professores por máquinas pode ameaçar a qualidade do ensino e a riqueza das interações humanas na academia. Assim, é fundamental que, mesmo com a adoção dessas tecnologias, o processo ensino-aprendizagem continue sendo humanizado e personalizado.

Vários estudos têm sido realizados para avaliar o impacto do uso do *ChatGPT* no ambiente acadêmico, como verificado por [60]. Essas pesquisas têm mostrado resultados positivos em termos de satisfação dos estudantes e eficiência na resolução de dúvidas. Os estudantes relatam que se sentem mais engajados e motivados quando interagem com o *ChatGPT*, e a rapidez das respostas contribui para uma melhor experiência de aprendizagem.

O uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico apresenta boas perspectivas. Além do *ChatGPT*, outras tecnologias, como a realidade virtual, podem ser integradas para criar ambientes de aprendizagem mais imersivos e interativos. Isso não apenas ajuda a prender os alunos na escola, aumentando seu engajamento e interesse, mas também demonstra outras formas de ensinar, permitindo que se explorem métodos que vão além do modelo tradicional de ensino. Essas abordagens inovadoras podem contribuir para uma educação mais dinâmica e adaptada às necessidades dos estudantes. Segundo [59], espera-se que o desenvolvimento de sistemas mais avançados sejam capazes de compreender e interagir de forma mais natural com os usuários. Essas inovações têm o potencial de transformar o ambiente acadêmico e melhorar ainda mais a experiência dos estudantes nesse espaço.

Inicialmente, salienta [61], a inteligência artificial era vista como um campo de estudo teórico, com poucas aplicações práticas. No entanto, com o desenvolvimento de algoritmos mais sofisticados e o aumento da capacidade computacional, a inteligência artificial passou a ser utilizada em diversas áreas, como reconhecimento de padrões, processamento de linguagem natural e aprendizado de máquina.

O uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico traz consigo uma série de benefícios. Primeiramente, essas ferramentas possibilitam o acesso a informações instantâneas, permitindo aos estudantes ampliarem seus conhecimentos de forma rápida e eficiente. Além disso, de acordo com [10], mostra-se que essas tecnologias permitem a personalização do ensino, adaptando-se às necessidades individuais dos alunos e proporcionando uma experiência de aprendizagem mais eficaz. Por fim, as ferramentas baseadas em inteligência artificial também auxiliam na resolução de problemas complexos, fornecendo insights e sugestões para a solução de desafios acadêmicos.

De acordo com [62], o uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico também apresenta desafios e preocupações. Um dos principais desafios é a falta de transparência nos algoritmos utilizados por essas ferramentas. Muitas vezes, os estudantes e professores não têm conhecimento sobre como esses sistemas funcionam e como são tomadas as decisões.

Diante do crescente uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico, torna-se necessário repensar a formação dos profissionais da área. De acordo com [2], é fundamental que os futuros professores e pesquisadores desenvolvam habilidades específicas para trabalhar com essas tecnologias, compreendendo seus princípios básicos e suas aplicações práticas. Além disso, é importante promover uma reflexão ética sobre o uso dessas ferramentas, discutindo questões relacionadas à privacidade, viés algorítmico e responsabilidade social.

Essas tecnologias têm o potencial de transformar positivamente a forma como aprendemos e ensinamos, proporcionando uma educação mais personalizada e eficiente. No entanto, de acordo com [63], ainda há desafios a serem superados, como a necessidade de desenvolver algoritmos mais transparentes e éticos, além de garantir a formação adequada dos profissionais da área. É fundamental que essas questões sejam abordadas de forma cuidadosa e responsável, visando o benefício de todos os envolvidos no processo educacional.

5.1 Inteligência artificial vs. inteligência humana

A divergência entre a inteligência humana e a inteligência artificial constitui um elemento essencial na compreensão das capacidades cognitivas distintas entre seres humanos e sistemas tecnológicos. Segundo [64], a inteligência humana é intrinsecamente caracterizada pela habilidade de compreensão, aprendizado e resolução de problemas. Este fenômeno complexo engloba uma diversidade de competências, tais como a capacidade de pensamento abstrato, resolução de problemas criativos e adaptação a contextos inéditos.

Contrastantemente, a inteligência artificial representa uma modalidade de inteligência simulada por meio de dispositivos mecânicos, como apontado por [64]. Esta esfera de pesquisa dinâmica destaca-se pela sua capacidade em evolução, potencialmente culminando na criação de máquinas com aptidões cognitivas comparáveis às humanas, tanto em termos de processamento mental quanto de ação.

Apesar das distinções fundamentais, identificam-se alguns pontos de convergência entre a inteligência humana e a artificial. Ambas envolvem a capacidade de processamento informacional e tomada de decisões, além de possuírem aplicabilidade na resolução de problemas e na aquisição de novos conhecimentos, como afirmado por [64].

A perspectiva revolucionária da inteligência artificial transcende as fronteiras do entendimento convencional sobre a inteligência, proporcionando uma abordagem que pode aprimorar nossa compreensão da cognição humana. Sua influência potencial sugere a possibilidade de conceber máquinas dotadas de uma inteligência e capacidade superiores às anteriormente imaginadas por [64].

5.2 Aprendizado com *ChatGPT*

A educação convencional caracteriza-se pela mediação do professor no processo de construção do conhecimento pelos alunos em sala de aula. Embora esse modelo tenha demonstrado eficácia, ele pode ser restritivo e inflexível, limitando a capacidade dos estudantes de progredir de acordo com seu ritmo e estilo de aprendizagem preferido. De acordo

com [65], a educação tradicional é muitas vezes geograficamente limitada, pois requer a presença física dos alunos na sala de aula.

A tecnologia surgiu como uma solução para abordar alguns desses desafios, possibilitando experiências de aprendizagem mais personalizadas e recursos envolventes e acessíveis. Plataformas de ensino online, livros digitais e ferramentas multimídia oferecem aos alunos acesso a uma ampla variedade de recursos, independentemente de sua localização geográfica, como afirmado por [65]. A tecnologia também facilita a colaboração e a comunicação entre alunos e professores, resultando em experiências de aprendizagem dinâmicas e interativas.

A análise das literaturas recentes mostra que a IA, especificamente o *ChatGPT*, vem sendo utilizada em diferentes contextos do processo ensino-aprendizagem com benefícios significativos. Entre as capacidades notáveis do *ChatGPT* está a criação de avaliações de aprendizagem, conforme exposto em [13], que relata a capacidade da IA de fornecer feedback em tempo real e relatórios detalhados, permitindo aos educadores reduzir a carga de trabalho em atividades administrativas e concentrar-se em iniciativas pedagógicas mais criativas.

Conforme [11], a inteligência artificial tem aplicabilidade em contextos educacionais diversificados. Eles demonstram que o *ChatGPT* pode ser instrumental na elaboração de materiais educacionais que não apenas envolvem os alunos, mas também promovem uma aprendizagem mais profunda e significativa. A capacidade da IA de fornecer feedback imediato e de adaptar-se às necessidades individuais dos alunos é um dos maiores avanços identificados, contribuindo para um ambiente de aprendizagem mais inclusivo e eficaz.

A criação de avaliações de aprendizagem é uma das áreas onde a IA tem demonstrado um impacto significativo. Segundo [13], a implementação do *ChatGPT* permite que os educadores desenvolvam avaliações com maior eficiência, utilizando feedbacks automatizados que são fornecidos aos alunos de maneira imediata. Essa funcionalidade não apenas reduz o tempo e esforço necessário para a criação de avaliações, mas também melhora a qualidade do ensino, permitindo que os educadores se concentrem em aspectos mais criativos e individualizados do processo educativo.

O *ChatGPT* tem sido destacado por sua capacidade de auxiliar na adaptação de materiais didáticos para diferentes públicos, segundo [11]. Esse estudo mostrou que os alunos que utilizam materiais ajustados pela IA, particularmente em disciplinas como ciências e história, apresentaram um maior engajamento e compreensão dos conteúdos, indicando um aumento na eficácia da construção do conhecimento.

O *ChatGPT* pode ser utilizado para analisar e resumir textos complexos, ajudando na economia de tempo para educadores e alunos, ao mesmo tempo em que melhora a compreensão dos conteúdos, conforme exposto em [13]. O *ChatGPT* é capaz de analisar a

entrada de linguagem natural dos alunos, o que permite aos sistemas de tutoria inteligentes compreenderem de forma mais eficaz suas perguntas, necessidades e preferências. Isso significa que os alunos podem receber respostas mais direcionadas e personalizadas para suas dúvidas, criando um ambiente de aprendizagem mais adaptado às suas necessidades individuais.

Segundo [66], um dos grandes problemas do uso da inteligência artificial está relacionado a dados falsos, uma vez que pesquisadores, estudantes e o público em geral podem criar dados plausíveis usando instruções simples para entrevistas ou modificar conjuntos de dados existentes de maneira bastante simples para criar um conjunto de dados que pareça plausível e seja então analisado e apresentado como se fosse coletado de indivíduos reais.

Capítulo 6

Benefícios dos *LLMs* (*ChatGPT*) no processo ensino-aprendizagem

Este capítulo explora o impacto da inteligência artificial, especialmente do *ChatGPT*, no ambiente acadêmico, destacando os benefícios que essa tecnologia pode trazer para esse contexto. São abordados tópicos como a criação de avaliações de aprendizagem 6.1, o aprimoramento da prática pedagógica 6.2 e o oferecimento de tutoria personalizada 6.3, a criação de esboços de artigos e pesquisas 6.4, o debate de ideias 6.5, a conscientização ética 6.6, a acesso à informação 6.7 e a redução de ansiedade 6.8. Com base na análise de diversos autores, discute-se como essas ferramentas podem otimizar processos educacionais, reduzir a carga de trabalho de educadores, personalizar o ensino e promover um processo ensino-aprendizagem mais eficiente e inclusivo. O capítulo ainda examina aspectos práticos e éticos do uso da IA no ambiente acadêmico, destacando seu potencial para transformar o modo como a construção do conhecimento é feita.

No século da informação e da revolução tecnológica, o cenário acadêmico enfrenta uma transformação sem precedentes impulsionada pela inteligência artificial. A integração dessas tecnologias inovadoras está redefinindo não apenas a maneira como aprendemos, mas também como os educadores ensinam e como as instituições educacionais operam. De acordo com [62] a crescente disponibilidade de ferramentas baseadas em IA oferece oportunidades sem igual para otimizar o processo de aprendizagem, personalizar o ambiente acadêmico e promover uma abordagem mais eficaz e inclusiva ao ensino. Neste contexto, esse capítulo tem o foco de apresentar os benefícios do uso dessas ferramentas do ponto de vista de diferentes autores.

Segundo [13], alguns dos possíveis benefícios que o uso da ferramenta *ChatGPT* poderia trazer dentro do contexto do ambiente acadêmico. A saber: criação de avaliações de aprendizagem (*creating learning assessment*), aprimoramento da prática pedagógica (*enhancing pedagogical practice*), oferecimento de tutoria personalizada (*offering virtual*

personal tutoring,), criação de esboços de artigos e pesquisas (*creating an essay or research article outline*) e debate de ideias (*brainstorming ideas*).

A pesquisa de [11] analisa o *ChatGPT* através do modelo *SWOT* (*Strengths, Weaknesses, Opportunities and Threats*), em tradução livre: forças, fraquezas, oportunidades e ameaças. A partir dele, obteve-se o benefício da conscientização ética.

O benefício do acesso à informação foi trazido por [14]. Trata-se de um artigo voltado para as aplicações práticas para o uso do *ChatGPT* e que corrobora com vários benefícios trazidos por outros textos, bem como traz os próprios.

Dentre todos os benefícios listados, um benefício bastante inusitado foi o trazido por [15]: a redução da ansiedade. Trata-se de um estudo qualitativo que focou a pesquisa em entrevistas junto às partes interessadas (*stakeholders*): alunos e professores.

6.1 Criação de avaliações de aprendizagem

Segundo [13], o *ChatGPT* fornece aos educadores meios inovadores e flexíveis para desenvolver avaliações de aprendizagem, incluindo *feedbacks* e relatórios em tempo real. Além disso, Encontra-se em [13] a discussão sobre o tempo gasto pelos professores na criação de mecanismos de avaliação e como o uso do *ChatGPT* permitiria aos professores reduzir sua carga de trabalho e focar na criação de planos de aula inovadores, desenvolvimento profissional e em oferecer suporte individualizado de *coaching* e mentoria aos alunos.

De acordo com [11], o *ChatGPT* pode desempenhar um papel ativo de assistência às faculdades na criação materiais educativos envolventes e inovadores. Eles ainda fornecem vários exemplos de conteúdos gerados para auxiliar no processo ensino-aprendizagem em diferentes áreas como turismo, sustentabilidade, ciência e física. Além disso, [11] salienta ainda que a maioria dos alunos preferem estudar materiais históricos modificados pelo *ChatGPT*.

Segundo [62], outro benefício do uso do *ChatGPT* é a capacidade de fornecer *feedback* imediato aos alunos. Ao realizar atividades ou responder perguntas, o sistema pode avaliar as respostas dos estudantes e fornecer comentários construtivos instantaneamente. Isso ocorre, como aponta [8], porque essas ferramentas são capazes de analisar o desempenho e as características de cada estudante, identificando suas dificuldades e oferecendo recursos específicos para auxiliá-los em seu processo de aprendizagem. Esse *feedback* contínuo é fundamental para o processo de aprendizagem, permitindo que os alunos corrijam seus erros e melhorem seu desempenho de forma mais eficiente.

6.2 Aprimoramento da prática pedagógica

Aprimorar a eficiência e a produtividade no meio acadêmico é de extrema importância para que estudantes e pesquisadores alcancem seus objetivos com excelência. A demanda por resultados de alta qualidade em um curto período de tempo tem levado à necessidade de utilizar ferramentas baseadas em inteligência artificial para otimizar os processos acadêmicos. Essas ferramentas, como defende [17], podem auxiliar na organização e gestão do tempo, permitindo que os acadêmicos se dediquem mais às atividades essenciais, utilizando o *ChatGPT* para tarefas repetitivas ou periféricas com menos importância.

De acordo com [13], são muitos os recursos inovadores fornecidos pelo *ChatGPT*. Dentre esses recursos, destaca-se a capacidade que o *ChatGPT* tem de incentivar a participação dos alunos nas aulas, favorecendo o engajamento e o protagonismo. O texto cita, ainda, que a assistência do *ChatGPT* permite a adoção de abordagens de ensino criativas como o método da aula invertida, que favorece o aprendizado independente do aluno.

Esse ponto de vista é compartilhado por [11] que defende que, além da sua capacidade para criar materiais pedagógicos diversos e relevantes, o *ChatGPT* permite a interação dos estudantes como interlocutores, motivando-os a explorar e aprofundar seus conhecimentos em tópicos de seu interesse, favorecendo a autoaprendizagem.

As oportunidades para o ensino personalizado gerados pelo *ChatGPT*, são um grande avanço para o aprimoramento da prática pedagógicas segundo [10]. No mesmo artigo é mostrado que, com essa ferramenta, cada estudante pode ter um acompanhamento mais individualizado e adaptado às suas necessidades específicas. Acredita-se que tais ferramentas poderiam identificar as lacunas de conhecimento de cada aluno e fornecer materiais e exercícios personalizados, contribuindo para um aprendizado mais eficiente. Dessa forma, é possível tornar o processo de ensino mais dinâmico e envolvente, facilitando a aprendizagem.

Segundo [59], uma das principais dificuldades enfrentadas pelos estudantes e pesquisadores no ambiente acadêmico que é a sobrecarga de informações. Com o avanço da tecnologia, o acesso a uma quantidade cada vez maior de dados tornou-se mais fácil, porém, isso também resultou em uma maior dificuldade em filtrar as informações relevantes. Além disso, de acordo com [59], a falta de tempo para realizar todas as tarefas necessárias é outra questão recorrente. Os acadêmicos muitas vezes se veem sob pressão para cumprir prazos apertados, o que pode comprometer a qualidade do trabalho produzido.

O uso de ferramentas baseadas em inteligência artificial, como o *ChatGPT*, pode ser um assistente para esses problemas. O *ChatGPT* possui funcionalidades que contribuem para a melhoria da eficiência e produtividade acadêmica. Por exemplo, sua capacidade de responder dúvidas rapidamente pode ajudar os estudantes a obterem respostas imediatas para suas perguntas, evitando assim perda de tempo na busca por informações.

Além disso, o *ChatGPT* pode fornecer sugestões relevantes para a pesquisa, auxiliando os pesquisadores na geração de ideias e no aprofundamento de seus estudos, como aponta [4].

Ao automatizar tarefas corriqueiras, o meio acadêmico abre espaço para o florescimento da pesquisa e do aprendizado, permitindo que indivíduos se dediquem a atividades de maior relevância. A realização de atividades rotineiras, como a coleta e organização de dados, a formatação de documentos e a busca por informações relevantes, pode consumir uma quantidade significativa de tempo que poderia ser direcionado para atividades mais complexas e criativas. Segundo [8], a utilização de ferramentas baseadas em inteligência artificial, como o *ChatGPT*, tem se mostrado um assistente eficiente para automatizar essas tarefas repetitivas.

Em relação à qualidade das pesquisas, essas ferramentas podem auxiliar na identificação e análise de informações relevantes, contribuindo para a produção de trabalhos mais completos e embasados. Ademais, ferramentas como o *ChatGPT* utilizam grandes bancos de dados, permitindo cruzamento entre informações de diferentes fontes e permitindo o acesso fácil através de uma interface amigável, com isso elas facilitam a comunicação entre os membros da comunidade acadêmica, permitindo o compartilhamento rápido e eficiente de conhecimentos e ideias. Esse uso, evidenciado em [67], promove a colaboração e o desenvolvimento de projetos conjuntos, enriquecendo o ambiente acadêmico como um todo.

De acordo com [63], a interação entre humanos e máquinas no ambiente acadêmico é essencial para promover um processo de ensino-aprendizagem mais completo. O *ChatGPT* pode complementar o trabalho dos professores ao oferecer suporte individualizado aos alunos, liberando tempo para que os educadores se concentrem em atividades mais complexas, como debates e projetos colaborativos. Dessa forma, a combinação da expertise humana com as capacidades da inteligência artificial pode potencializar a qualidade do ensino.

Segundo [4], as ferramentas baseadas em IA possuem a capacidade de analisar grandes volumes de dados educacionais, identificando padrões e tendências que podem ser utilizados para melhorar o planejamento pedagógico. Com base nessas análises, é possível tomar decisões mais embasadas e eficientes, direcionando recursos e estratégias de ensino de forma mais precisa. Portanto, o uso de ferramentas de IA no ambiente acadêmico permite um planejamento pedagógico mais informado e eficiente, otimizando a alocação de recursos e estratégias de ensino.

Outro benefício do uso de ferramentas baseadas em IA no ambiente acadêmico, como abordado em [58], é a possibilidade de automação de tarefas administrativas. Essas ferramentas podem realizar atividades rotineiras, como correção automática de provas e

avaliações, agendamento de horários e organização de dados. Com isso, os profissionais da área educacional podem economizar tempo e se dedicar a atividades mais estratégicas e de maior valor agregado, como o planejamento pedagógico e a orientação dos alunos.

Segundo [60], possíveis avanços tecnológicos podem incluir melhorias na compreensão do contexto das perguntas e respostas, bem como a incorporação de técnicas de análise de sentimentos em textos acadêmicos, que consiste em realizar classificação de textos em linguagem natural, com o propósito de saber o sentimento expresso pelo texto. Esses avanços podem beneficiar áreas como a psicologia, sociologia e ciências humanas, permitindo uma compreensão mais profunda das emoções e opiniões expressas nos textos acadêmicos.

6.3 Oferecimento de tutoria personalizada

Conforme apontado por [13], o *ChatGPT* pode agir como um tutor virtual, oferecendo suporte individualizado a estudantes de acordo com suas necessidades específicas. Esse sistema de tutoria, ainda de acordo com [13], é diferente dos métodos tradicionais, pois está disponível a qualquer hora, oferecendo feedback imediato e respostas personalizadas a perguntas dos alunos, o que pode ser particularmente útil para alunos que precisam de ajuda extra fora do horário escolar ou para aqueles que preferem aprender em um ambiente mais privado, porém sem descartar a orientação dos professores.

A personalização do ensino e aprendizagem no ambiente acadêmico, como apresentado por [61], tem uma grande importância para assegurar a eficácia do processo educacional e o envolvimento dos discentes. Cada estudante possui características individuais, como ritmo de aprendizado, interesses e necessidades específicas, que devem ser consideradas para promover a construção do conhecimento de forma efetiva. A abordagem tradicional de ensino, baseada em aulas expositivas e padronizadas, muitas vezes não consegue atender às demandas individuais dos alunos, resultando em desmotivação e dificuldades de aprendizagem.

De acordo com [58], as limitações do ensino tradicional são evidentes quando se trata da personalização do ensino. Os professores têm um número limitado de horas para ministrar as aulas e precisam seguir um currículo pré-estabelecido. Isso impede que eles dediquem tempo suficiente para atender às necessidades individuais de cada aluno. No entanto, o uso de ferramentas baseadas em inteligência artificial, como o *ChatGPT*, pode superar essas limitações ao oferecer uma abordagem mais adaptativa e flexível.

Uma das principais consequências do uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico é o estímulo à autonomia dos estudantes, como abordado em [9]. Ao ter acesso facilitado a informações e conhecimentos, os estudantes são incentivados a buscar ativamente seu próprio aprendizado e desenvolver habilidades de pesquisa

independentes. Além disso, essas ferramentas aumentam a eficiência na realização de tarefas acadêmicas, permitindo que os estudantes economizem tempo e esforço na busca por informações.

Como apresentado em [59], uma das principais vantagens do uso de ferramentas como o *ChatGPT* na personalização do ensino é a capacidade de adaptar-se ao ritmo de aprendizado de cada aluno. O *ChatGPT* pode identificar as lacunas no conhecimento do estudante e fornecer explicações adicionais ou exercícios extras para fortalecer essas áreas específicas. Dessa forma, cada aluno pode avançar em seu próprio ritmo, garantindo uma compreensão mais sólida dos conteúdos.

Além disso, o *ChatGPT* pode ser utilizado como um tutor virtual, capaz de fornecer explicações e tirar dúvidas dos alunos. Essa interação individualizada permite que os estudantes tenham acesso a um suporte constante, mesmo fora do horário das aulas. O *ChatGPT* pode oferecer respostas imediatas e precisas, segundo [59], auxiliando os alunos na resolução de problemas e no desenvolvimento de suas habilidades cognitivas.

6.4 Criação de esboços de artigos e pesquisas

Dentre as funcionalidades destacadas para pesquisadores e estudantes, assume destaque, conforme apontado por [13], sua capacidade de gerar esboços para artigos e pesquisas. Embora haja certa discordância acerca do uso de inteligências artificiais generativas para a geração do conteúdo dos artigos, segundo [13], o *ChatGPT* pode ser de grande ajuda para organizar ideias e ajudar a estruturar os textos, apesar da possível necessidade de ajustes para garantir sua utilização.

A elaboração de trabalhos acadêmicos de qualidade é de extrema importância no ambiente acadêmico, uma vez que contribui para o desenvolvimento do conhecimento científico e para a formação dos estudantes. Nesse sentido, pode ser útil utilizar ferramentas baseadas em inteligência artificial para auxiliar nesse processo. De acordo com [18], essas ferramentas podem fornecer sugestões de conteúdo, corrigir erros gramaticais e auxiliar na estruturação do texto, o que contribui para a produção de trabalhos mais consistentes e bem fundamentados.

Em concordância com [18], uma das principais vantagens apontada por [10] é a melhoria da eficiência e produtividade dos estudantes ao obterem respostas rápidas e precisas para suas dúvidas. Isso permite que eles avancem em seus estudos de forma mais ágil, sem perder tempo buscando informações em diferentes fontes. Além desse ponto [10] ainda expõe que o *ChatGPT* pode auxiliar os estudantes na elaboração de trabalhos acadêmicos, fornecendo sugestões e correções gramaticais.

Ainda segundo [3], outro impacto positivo que o uso do *ChatGPT* traz é a ampliação das possibilidades de pesquisa, já que essa ferramenta pode auxiliar na busca por referências bibliográficas e na identificação de novas abordagens para determinado tema.

Conforme apresentado por [63], e também por [67], lidar com grandes volumes de informações é um grande desafio para os pesquisadores e que o uso da inteligência artificial pode ajudar na análise desses dados por meio da identificação de padrões e tendências que seriam difíceis ou demorados para serem identificados manualmente.

6.5 Debate de ideias

Seguindo um caminho semelhante ao da criação de esboços para artigos e pesquisas, segundo [13], a possibilidade de usar o *ChatGPT* para o aprimoramento de ideias. São citados, inclusive, artigos inteiros produzidos pela ferramenta de inteligência artificial, tendo sido necessários apenas alguns ajustes de edição e reorganização.

Ferramentas baseadas em inteligência artificial, como o *ChatGPT*, têm o potencial de aprimorar significativamente o acesso à informação e ao conhecimento no contexto acadêmico. Essas tecnologias, como aponta [67], podem processar grandes volumes de dados em tempo real, oferecendo respostas precisas e relevantes para as consultas dos usuários. Como resultado, os estudantes podem obter as informações necessárias de forma ágil, sem a necessidade de investir horas em pesquisas em múltiplas fontes.

Com a ajuda do *ChatGPT*, os estudantes e pesquisadores podem obter informações mais precisas e atualizadas, o que contribui para a produção de trabalhos de maior relevância, contribuindo com a melhora da qualidade das produções científicas. Conforme exposto por [61], o uso dessa ferramenta pode reduzir o estresse causado pela pressão por prazos apertados, uma vez que ela agiliza o processo de busca por informações e geração de conteúdo viabilizando assim a produção de trabalhos mais relevantes.

De acordo com [17], uma das principais vantagens é a agilidade na busca por informações. Com apenas algumas perguntas, os estudantes podem acessar uma vasta quantidade de recursos relevantes para seus estudos. Além disso, essas ferramentas

O *ChatGPT* possui diversas aplicações potenciais no ambiente acadêmico, como apresentado por [59]. Ele pode ser utilizado para auxiliar na pesquisa bibliográfica, fornecendo sugestões de artigos relevantes com base nas palavras-chave inseridas pelo usuário. A ferramenta também pode responder dúvidas específicas dos estudantes, oferecendo explicações detalhadas sobre conceitos complexos.

6.6 Conscientização ética

De acordo com [11] defende que um dos papéis das faculdades é elevar a conscientização ética entre seus estudantes. Neste sentido, os autores consideram que o *ChatGPT* pode engajar os estudantes em discussões sobre as implicações éticas do uso de inteligências artificiais, cultivando uma implementação responsável da inteligência artificial.

Segundo [68], o aprimoramento do raciocínio crítico no contexto acadêmico é fundamental para preparar profissionais com a capacidade de enfrentar os complexos desafios impostos pela era da inteligência artificial, permitindo uma abordagem mais ética e consciente. Ele ainda mostra que essas habilidades são fundamentais para que os estudantes sejam capazes de analisar, avaliar e interpretar informações de forma objetiva e embasada, além de serem essenciais para o processo de tomada de decisão e resolução de problemas complexos. O pensamento crítico permite aos estudantes questionarem suposições, identificarem falácias e argumentarem com base em evidências sólidas, contribuindo para um ambiente acadêmico mais rico e produtivo.

Para maximizar os benefícios do uso do *ChatGPT* no desenvolvimento das habilidades de pensamento crítico, é importante adotar estratégias adequadas, como aponta [10]. Os estudantes devem ser incentivados a refletir sobre as respostas geradas pela ferramenta, questionando-as e comparando-as com outras fontes confiáveis. Outro ponto levantado foi que o debate e a colaboração entre os estudantes podem ser estimulados como forma de promover uma análise mais profunda das questões abordadas pelo *ChatGPT*. Dessa forma, os estudantes podem desenvolver uma postura crítica em relação às informações fornecidas pela ferramenta e aprimorar suas habilidades de pensamento crítico. Além disso, o *ChatGPT* pode estimular a curiosidade intelectual dos estudantes, incentivando-os a buscar conhecimento além do que é fornecido pela ferramenta [3].

6.7 Acesso à informação

A facilitação do acesso à informação e ao conhecimento é importante para o sucesso individual e coletivo no ambiente acadêmico. A disponibilidade de recursos e materiais relevantes desempenha um papel central no desenvolvimento do ensino e da pesquisa. Conforme apresentado por [60], a capacidade de acessar informações atualizadas de forma rápida e fácil permite que os estudantes acompanhem os avanços em suas áreas de estudo. Além disso, o acesso facilitado promove a democratização do conhecimento, garantindo que estudantes de diversas origens possam usufruir das mesmas oportunidades educacionais.

De acordo com [14], o *ChatGPT* tem potencial para atuar como uma ferramenta de pesquisa, ajudando os alunos a localizar e organizar dados para trabalhos e outros tipos de estudo, facilitando o acesso a fontes de informação relevantes. A esse potencial soma-se, ainda segundo [14], sua capacidade de fazer traduções e processar comandos complexos, tudo isso de forma rápida, permitindo acesso eficiente à informação. A pesquisa feita por [3] corrobora àquela feita por [14], uma vez que defende que um dos principais benefícios do uso do *ChatGPT* no ambiente acadêmico é a agilidade na obtenção de respostas para dúvidas. Com essa ferramenta, os estudantes podem ter acesso rápido a informações relevantes, facilitando o processo de pesquisa e estudo.

Seguindo o mesmo raciocínio, o exposto em [59] defende que a democratização do acesso ao conhecimento representa um impacto importante das ferramentas de inteligência artificial. Segundo essa autora, estas tecnologias têm o potencial de reduzir significativamente a exclusão digital, permitindo que estudantes de diferentes níveis socioeconômicos acessem informações e recursos educativos valiosos. A inclusão digital proporcionada por essas ferramentas não apenas promove maior igualdade de oportunidades, mas também possibilita que alunos de diversas origens tenham acesso a uma construção de conhecimento de qualidade. Dessa forma, a inteligência artificial pode contribuir para a construção de um ambiente educacional mais inclusivo e acessível.

6.8 Redução da ansiedade

Segundo entrevistas realizadas por [15], os estudantes relatam que provas e "tarefas de casa" os fazem sentir ansiosos. Neste cenário, os autores defendem que o *ChatGPT* entra como um recurso de fácil acesso e sem o receio de críticas ou avaliações depreciativas, impactando positivamente na experiência educacional dos estudantes.

Outro exemplo que pode demonstrar o potencial para redução da ansiedade foi relatado por [18]. Os autores citam um estudo japonês que contou com um *chatbot* para simular uma conversa em um restaurante. Segundo [18], após a interação com o *chatbot*, os resultados mostraram que a confiança e o desejo de se comunicar aumentaram, enquanto o nervosismo diminuiu.

Capítulo 7

Riscos dos *LLMs* (*ChatGPT*) no processo ensino-aprendizagem

Nesse capítulo é apresentado os riscos e preocupações éticas relacionados ao uso do *ChatGPT* no ambiente acadêmico. São discutidos problemas como a integridade acadêmica 7.1, avaliações de aprendizagem injustas 7.2, disseminação de informações imprecisas 7.3, perpetuação de preconceitos 7.4, dependência excessiva da IA 7.5 e riscos à privacidade 7.6. É discutido como o uso indiscriminado dessa tecnologia pode impactar negativamente o desenvolvimento intelectual e ético no ambiente acadêmico, além de exaltar a necessidade de regulamentação e práticas educacionais responsáveis.

Assim como são trazidos benefícios do uso do *ChatGPT*, a pesquisa de [13] lista, também, riscos associados ao seu uso em pesquisa e no ambiente acadêmico. As principais áreas de preocupação citadas são: riscos à integridade acadêmica (*Academic integrity issues*) 7.1, avaliações de aprendizagem injustas (*Unfair learning assessment*) 7.2, informação imprecisa (*Inaccurate information*) 7.3 e dependência excessiva da IA (*Over-reliance on AI*) 7.5.

Os riscos da perpetuação de preconceitos e àqueles relacionados à privacidade 7.6 foram introduzidos por [14]. Problemas esses que viriam a ser citados posteriormente por [16] e [18].

7.1 Riscos à integridade acadêmica

As preocupações éticas relacionadas ao uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico são importantes, visto que, com o uso dessas ferramentas, o aumento do plágio, um problema persistente ao longo dos anos, pode aumentar significativamente. Segundo [18], há a preocupação de que o uso dessas ferramentas possa ser

considerado plágio tendo em vista que a ferramenta usa textos de outros autores, sem mencioná-los de maneira apropriada.

Outra preocupação está relacionada à possibilidade de substituição dos profissionais humanos. De acordo com [69], o uso do *ChatGPT* pode levar à diminuição da demanda por professores e pesquisadores, uma vez que parte das tarefas acadêmicas poderia ser realizada por essa ferramenta. Para mitigar esse problema, é fundamental estabelecer diretrizes éticas para o uso do *ChatGPT* no ambiente acadêmico, garantindo que ele seja utilizado como uma ferramenta complementar ao trabalho humano e não como um substituto.

De acordo com [13], o *ChatGPT* pode ser usado indevidamente por estudantes para produzir respostas ou artigos sem declaração ou citação apropriada, levando a casos de plágio e direitos autorais. Isso pode resultar em má conduta acadêmica e desonestidade, prejudicando o sucesso acadêmico dos estudantes a longo prazo.

Além das preocupações já mencionadas, a falta de interação humana no processo de aprendizagem é outro aspecto importante a ser considerado. Como abordado em [69], o uso do *ChatGPT* não deve substituir completamente o diálogo e a troca de ideias entre estudantes e professores, elementos fundamentais para um aprendizado efetivo.

7.2 Avaliações de aprendizagem injustas

Embora a capacidade do *ChatGPT* de gerar textos seja amplamente reconhecida, conforme apontado por [13], tais textos podem ser, por vezes, difíceis de distinguir daqueles produzidos por humanos. Sendo assim, os autores levantam a possibilidade da terceirização do trabalho por parte dos estudantes, o que configuraria, segundo o próprio [13], uma vantagem injusta sobre outros estudantes que não fizeram o mesmo uso da ferramenta.

As conclusões formadas por [11] parecem concordar com a posição de [13] ao elencar a fraqueza *cheating and gaining an advantage in exams* (em português: trapacear e ganhar vantagem nos exames). Segundo [11], a capacidade do *ChatGPT* de passar exames (uma lista desses exames pode ser acessada na tabela 4.1 na seção 4.5) levanta preocupações significantes acerca da possibilidade do uso dessa ferramenta por estudantes para obter vantagens injustas nas avaliações de desempenho.

Partindo de um ponto de vista mais objetivo, segundo [14], uma das preocupações mais comuns no setor educacional é a de que os estudantes podem utilizar ferramentas como o *ChatGPT* para finalizar os "deveres de casa" (*homeworks*) e, então, "copiar e colar" (*copy and paste*) os resultados sem qualquer controle dos professores.

7.3 Informação imprecisa

Segundo [13], o *ChatGPT* pode produzir informações factualmente incorretas, como artigos inventados ou respostas imprecisas, o que pode causar efeitos prejudiciais quando essa informação é usada por estudantes ou professores em seus trabalhos.

Outro desafio é a necessidade de capacitação adequada para utilizá-las corretamente, pois a resposta gerada depende da qualidade das perguntas. É fundamental que os usuários compreendam as limitações dessas ferramentas e saibam interpretar corretamente os resultados gerados por elas, além de ter em mente uma intencionalidade na busca e conhecimento prévio sobre o assunto para poder fazer uma crítica do texto retornado.

Sendo assim, uma das principais preocupações é a possibilidade de disseminação de informações falsas ou enviesadas. Como o *ChatGPT* gera respostas com base em modelos pré-treinados, ele pode reproduzir informações imprecisas ou tendenciosas presentes nos dados utilizados para seu treinamento, como aponta [69]. Além disso, existe o risco de os estudantes se tornarem excessivamente dependentes da tecnologia, deixando de desenvolver suas próprias habilidades de pesquisa e pensamento crítico.

7.4 Perpetuação de preconceitos

De acordo com [2], é importante notar que esses sistemas podem apresentar erros ou vieses nos resultados, o que pode comprometer a qualidade das informações obtidas. Essa ideia parece concordar com [14], que reafirma esse perigo, destacando que qualquer viés utilizado no conjunto de treinamento da ferramenta, reflete em suas respostas. Ainda conforme [14], embora existam salvaguardas e mecanismos para impedir abusos na interação, a ferramenta ainda incorre no risco de gerar conteúdos injustos, racistas e sexistas.

A possibilidade de vieses e preconceitos serem reproduzidos pelas ferramentas baseadas em inteligência artificial, como o *ChatGPT*, no ambiente acadêmico é uma preocupação relevante. Essas ferramentas são alimentadas com grandes quantidades de dados, que podem conter vieses implícitos presentes na sociedade. Segundo [4], é possível que o *ChatGPT* reproduza esses vieses ao responder perguntas ou fornecer informações no contexto acadêmico. Isso pode resultar em respostas tendenciosas ou incompletas, prejudicando a qualidade do ensino e a formação dos estudantes. Um dos principais receios, segundo [18], é a substituição de profissionais por máquinas, o que poderia gerar vieses algorítmicos, gerando um processo enviesado sem uma pluralidade de ideais.

A falta de interação humana e personalização do ensino é outra questão que pode surgir com o uso excessivo de ferramentas baseadas em inteligência artificial no ambiente acadêmico, privando a exposição de diferentes ideias na construção do ensino-aprendizagem.

Essas ferramentas, abordados por [61], geralmente oferecem respostas padronizadas e não levam em consideração as características individuais dos estudantes. Isso pode resultar em uma experiência de aprendizagem menos enriquecedora, pois os estudantes podem sentir falta do contato humano e da possibilidade de discutir ideias e tirar dúvidas com um professor, tendo um ensino enviesado de acordo com o *prompt* do estudante.

Não obstante, de acordo com [13], pesquisadores deveriam passar a utilizar o *ChatGPT* para obter um melhor entendimento de suas vantagens e desvantagens. Os autores salientam, entretanto, que o uso deveria se restringir ao debate de ideias (em inglês, *brainstorming*) e criação de esboços, para prevenir falhas intencionais e não intencionais, além de fazer uma checagem rigorosa da precisão das informações fornecidas pela ferramenta.

7.5 Dependência excessiva da IA

Considerando todas os benefícios apontados até aqui, bem como as capacidades do *ChatGPT*, é razoável considerá-lo uma solução para diversos problemas enfrentados no dia a dia. É nesse sentido que [13] chama atenção para o risco de que os usuários possam se tornar excessivamente dependentes do *ChatGPT* para a geração de conteúdo, o que pode diminuir a criatividade e as habilidades de pensamento crítico dos estudantes.

A utilização de ferramentas baseadas em IA, como o *ChatGPT*, para auxiliar na produção de conteúdo acadêmico pode levar à diminuição da criatividade e originalidade dos trabalhos, habilidades essenciais acadêmicas. Essas ferramentas são capazes de gerar textos automaticamente, o que pode resultar em uma padronização dos trabalhos e na perda da autenticidade do pensamento dos estudantes, como apontado por [10]. A produção de conteúdo acadêmico requer a capacidade de pensar criticamente e desenvolver ideias originais, habilidades que podem ser comprometidas com o uso excessivo dessas ferramentas, deixando tanto docentes como discentes dependentes da mesma, como aponta [18].

De acordo com [59], a resolução de problemas complexos, a análise crítica de informações e a tomada de decisões são competências essenciais para o sucesso acadêmico e profissional. No entanto, ao depender exclusivamente das respostas geradas por essas ferramentas, os estudantes podem deixar de exercitar essas habilidades, comprometendo seu desenvolvimento. A conclusão de [18] vai além e ressalta que existe o risco de os usuários se tornarem excessivamente dependentes dessa ferramenta para a realização de tarefas acadêmicas, deixando de desenvolver habilidades essenciais, como a capacidade de pesquisa e análise crítica.

Segundo [4], os estudantes enfrentam diversas dificuldades na elaboração de trabalhos acadêmicos. A falta de conhecimento técnico é uma das barreiras mais significativas, pois muitos alunos carecem de habilidades avançadas em pesquisa e redação. Além disso, a

falta de tempo agrava a situação, já que os estudantes precisam dividir sua atenção entre os estudos e outras atividades. Outro desafio frequente é a organização das ideias, com muitos alunos lutando para estruturar seus textos de forma coerente e convincente.

É importante considerar, conforme expresso em [9], que o uso excessivo dessas ferramentas pode representar uma ameaça à autonomia intelectual dos estudantes pois, assim como explicado em [17], a simples digitação de uma pergunta em um computador não fará com que competências como pensar criticamente e identificar

7.6 Riscos à privacidade

É importante ressaltar que o uso de ferramentas baseadas em IA no ambiente acadêmico também levanta preocupações éticas, como apontado em [3]. A privacidade dos dados dos alunos é uma questão fundamental nesse contexto, uma vez que essas ferramentas coletam e armazenam informações pessoais dos estudantes, existindo, assim, a possibilidade de vieses algorítmicos, ou seja, a reprodução de desigualdades sociais por meio do uso dessas ferramentas e a análise indevida dos dados desses estudantes. É necessário garantir que os algoritmos sejam imparciais e não perpetuem preconceitos ou discriminações.

Faz-se necessário, portanto, que os alunos e membros da equipe docente estejam cientes dos riscos à privacidade apresentados pelo *ChatGPT*. Pelo menos, é o que apontam [14]. É necessário garantir a privacidade dos alunos ao lidar com dados pessoais sensíveis durante a interação com o *ChatGPT*. A transparência e a responsabilidade na utilização dessas ferramentas são fundamentais para garantir um processo de ensino-aprendizagem mais justo e inclusiva de acordo com [1].

Segundo [68], existem dois grandes problemas: o primeiro é o viés algorítmico 3.3 que é um problema recorrente nessas tecnologias, podendo resultar em discriminação ou exclusão de determinados grupos. O segundo é a privacidade dos dados dos estudantes, também uma questão relevante, uma vez que o uso dessas ferramentas implica na coleta e armazenamento de informações pessoais, o que pode gerar preocupações em relação à segurança e uso indevido desses dados.

Capítulo 8

Recomendações dos *LLMs* (*ChatGPT*) no processo ensino-aprendizagem

Este capítulo tem como objetivo reunir as recomendações dadas por cada um dos autores para reduzir os riscos observados pelos mesmos. Este capítulo divide suas seções de acordo com os riscos apresentados no capítulo anterior. Essa divisão busca facilitar a relação entre os riscos e suas respectivas recomendações.

Algumas das recomendações propostas possuem grande abrangência de impacto em dois ou mais riscos, como, por exemplo, a proposta feita por [7] de estabelecer regras claras para utilização de ferramentas baseadas em inteligência artificial. Nesses casos, buscou-se relacionar a recomendação ao risco mais afetado por ela.

8.1 Riscos à integridade acadêmica

Conforme explorado anteriormente, riscos desta categoria estão relacionados principalmente à ocorrência do plágio. Como forma de evitar esses riscos, o [7] salienta que o *ChatGPT* geralmente produz referências bibliográficas fictícias, bastando uma simples conferência de fontes para encontrá-las. Não obstante, o autor também considera a utilização de ferramentas identificadoras de plágio baseadas em inteligência artificial, uma vez que detectores convencionais podem não ser suficientes neste caso.

A recomendação de [7] é que as instituições estabeleçam regras claras para a utilização de ferramentas como o *ChatGPT* e que se forneçam capacitações para docentes e discentes, de forma a facilitar a identificação de trabalhos realizados com a ajuda da ferramenta, enfatizando a necessidade de se utilizar fontes confiáveis e as consequências da má conduta acadêmica. Além do estabelecimento de regras, os autores de [11] consideram que a

implementação de sistemas robustos de monitoramento para rastrear as interações dos alunos com o *ChatGPT* pode ajudar a identificar possíveis situações de plágio.

8.2 Avaliações de aprendizado injustas

Como forma de evitar o risco das avaliações de aprendizagem injusta, a recomendação de [13] é considerar a aplicação de avaliações formativas em detrimento das somativas. De acordo com [70], a avaliação formativa possui o foco no aprendizado durante avaliação, em contraste à avaliação somativa, que avalia o resultado do que foi aprendido ao longo de todo o programa educacional. Ainda segundo [70], a avaliação somativa costuma ter valores e unidades predefinidos e comparáveis, enquanto a formativa utiliza atividades mais focadas na troca de conhecimentos entre professores e alunos.

As conclusões de [7] parecem ser semelhantes às proposta por [13] ao recomendar a utilização de questões que estimulem a análise e criatividade em detrimento de provas que exijam mera memorização. O autor ainda recomenda a utilização de exames totalmente livres de recursos digitais, bem como a incorporação de recursos midiáticos como músicas, vídeos e imagens.

Em [15] retoma-se, também, a importância de se projetar avaliações que exijam pensamento crítico, resolução de problemas e expressão criativa, uma vez que, segundo os autores, são tarefas que a inteligência artificial não consegue replicar facilmente.

8.3 Informação imprecisa

De acordo com [60], é importante considerar algumas preocupações relacionadas ao uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico. Uma delas é a confiabilidade das informações fornecidas pelo *ChatGPT*. Embora essa ferramenta seja capaz de gerar respostas rápidas, nem sempre essas respostas são precisas ou corretas. Portanto, é necessário desenvolver mecanismos de verificação das informações geradas pelo *ChatGPT*, garantindo assim a confiabilidade dos resultados obtidos, o que compromete a qualidade do conhecimento transmitido aos estudantes.

Para evitar o risco de incluir informações imprecisas, a recomendação feita por [13] é incentivar os estudantes a verificar, analisar e editar as respostas geradas pelo *ChatGPT* para garantir sua precisão, familiarizando-se com as forças e limitações da ferramenta. Algo semelhante ao que propõe [7] ao propor ensinar os estudantes sobre a importância de usar fontes autoritativas. Em concordância com [13], é sugerido por [15] que os estudantes cruzem as informações geradas por inteligência artificial com fontes acadêmicas confiáveis para promover o hábito de discernimento e reforçar a importância da precisão.

8.4 Dependência excessiva da IA

A fim de impedir que os estudantes se tornem excessivamente dependentes de ferramentas baseadas em inteligência artificial, em [11] faz-se a recomendação de possuir um robusto sistema de monitoramento do uso de tais ferramentas, além de defenderem que os educadores devem conceber tarefas e atividades que promovam o pensamento crítico e a criatividade, garantindo que os alunos se envolvam com o material, em vez de dependerem apenas de respostas geradas pela inteligência artificial.

Segundo [15], é importante que se oriente os alunos a equilibrarem o uso da assistência de ferramentas baseadas em inteligência artificial com o desenvolvimento de suas próprias habilidades, enfatizando a importância da colaboração com colegas e das interações com o corpo docente para o crescimento intelectual. A proposta de [16] se mostra semelhante à de [15], defendendo uma abordagem construtivista que enfatize a importância da produção de conhecimento pessoal juntamente com a assistência da inteligência artificial.

8.5 Perpetuação de preconceitos

Segundo [14], uma recomendação eficaz para reduzir o risco da perpetuação de preconceitos é a incorporação de feedback humano no processo de treinamento do modelo. Isso pode melhorar significativamente as respostas geradas, garantindo que elas estejam mais alinhadas com a precisão humana e os padrões de linguagem natural, o que ajuda a otimizar o sistema. Além disso, é fundamental diversificar os dados de treinamento utilizados para o *ChatGPT*. Ao incluir uma gama mais ampla de perspectivas e experiências no conjunto de dados, os desenvolvedores podem mitigar o risco de viés que pode surgir de um conjunto de dados limitado ou tendencioso.

Outra abordagem importante, ainda de acordo com [14], é a monitoração e avaliação contínuas das saídas do modelo. Isso pode ajudar a identificar e corrigir quaisquer viés que possa surgir ao longo do tempo, garantindo que a inteligência artificial permaneça uma ferramenta justa e eficaz para os usuários. Com a implementação dessas recomendações, é possível reduzir os problemas de viés e melhorar a confiabilidade do ChatGPT em contextos educacionais.

8.6 Riscos à privacidade

Para mitigar os riscos à privacidade, a recomendação de [14] é implementar a anonimização dos dados, garantindo que as informações pessoais de alunos e educadores sejam protegidas durante o processo de treinamento do modelo. Além disso, a adoção de me-

didadas robustas de segurança de dados, como criptografia e protocolos de comunicação seguros, é essencial para evitar acessos não autorizados e proteger as interações com a inteligência artificial. A conscientização e a capacitação dos usuários sobre os riscos associados ao uso de ferramentas de inteligência artificial também são cruciais, pois ajudam a garantir que informações sensíveis não sejam compartilhadas inadvertidamente durante as interações com o *ChatGPT*.

Outra recomendação importante, segundo [14], é o estabelecimento de diretrizes e políticas claras para o uso de ferramentas baseadas em inteligência artificial nas instituições educacionais. Essas diretrizes devem assegurar a conformidade com as leis de privacidade e os padrões éticos, definindo casos de uso aceitáveis e as responsabilidades de educadores e alunos ao utilizar o *ChatGPT*. Além disso, a monitorização contínua e a avaliação do desempenho da IA podem ajudar a identificar vulnerabilidades e áreas que necessitam de melhorias, permitindo ajustes oportunos nas práticas e políticas de proteção de dados.

Capítulo 9

Resultados

Segundo [59], a incorporação de tecnologias, como o *ChatGPT*, no ambiente educacional está promovendo uma transformação na forma como o conhecimento é contruído. Este capítulo explora o papel do *ChatGPT* como uma ferramenta de apoio ao ensino e à aprendizagem, analisando seus benefícios e os riscos associados, apresentados de forma resumida em tabelas. Nessas tabelas, os pontos positivos e negativos são agrupados com base nas contribuições de diferentes autores, permitindo uma visão comparativa e estruturada. Por fim, são apresentadas propostas para um uso ético e equilibrado da tecnologia, garantindo que ela complemente o ensino tradicional e contribua para o aprimoramento da prática pedagógica e da formação dos estudantes.

O processo de ensino-aprendizagem hoje, afirma [59], está saindo dos moldes tradicionais, onde o professor estava no centro do conhecimento, para um modelo onde o conhecimento é construído de forma conjunta, abrindo o leque do aprendizado, melhorando a expressão autônoma e individualista dos indivíduos. Nesse contexto é necessário uma abordagem educacional em que se tenha ativamente no processo de ensino-aprendizagem o recurso de tecnologias, para assim o conhecimento fluir dentro da nossa sociedade, proporcionando uma reestruturação do ensino e o desenvolvimento de novos ambientes de aprendizado.

O *ChatGPT* tem se mostrado uma ferramenta poderosa para o apoio ao ensino e à aprendizagem, proporcionando aos estudantes uma maneira mais acessível e personalizada de explorar conteúdos acadêmicos. De acordo com [14], um dos principais benefícios do *ChatGPT* é sua capacidade de auxiliar na compreensão de conceitos complexos, especialmente em disciplinas que exigem um alto nível de abstração. Essa ferramenta permite que os alunos façam perguntas e obtenham respostas detalhadas e adaptadas ao seu nível de entendimento, algo que nem sempre é possível em uma sala de aula tradicional com um número elevado de estudantes. A tabela 9.1 apresenta a compilação dos benefícios e riscos encontrados ao longo das pesquisas realizadas para produção deste trabalho.

Tabela 9.1: Tabela contendo o resumo dos benefícios e riscos apontados pelos autores das publicações estudadas.

Benefícios (Cap. 6)	Riscos (Cap. 7)
Criação de avaliações de aprendizagem 6.1	Riscos a integridade acadêmica 7.1
Aprimoramento da prática pedagógica 6.2	Avaliações de aprendizagem injustas 7.2
Oferecimento de tutoria personalizada 6.3	Informação imprecisa 7.3
Criação de esboços de artigos e pesquisas 6.4	Perpetuação de preconceitos 7.4
Debate de ideias 6.5	Dependência excessiva da IA 7.5
Conscientização ética 6.6	Riscos à privacidade 7.6
Acesso à informação 6.7	
Redução da ansiedade 6.8	

A análise de publicações relacionadas ao uso do *ChatGPT* no processo de ensino-aprendizagem permitiu elencar diversos pontos positivos (benefícios) e diversos pontos negativos (riscos). A tabela 9.2 contém uma sumarização desses pontos positivos e negativos a partir da visão de cada um dos autores. É importante destacar que a tabela não contém uma lista exaustiva de todos os pontos trazidos pelos autores. Muitos pontos foram omitidos por não representar uma mudança direta no processo ensino-aprendizagem, enquanto outros resultaram de pequenas adaptações em seus títulos ou generalizações de seus conceitos para permitir uma melhor conversação entre autores.

Tabela 9.2: Tabela contendo os benefícios e riscos segundo cada um dos autores estudados para este trabalho

Autor	Benefícios	Riscos
Sok et al.[13]	<p>Criação de avaliações de aprendizagem 6.1</p> <p>Aprimoramento da prática pedagógica 6.2</p> <p>Oferecimento de tutoria personalizada 6.3</p> <p>Criação de esboços de artigos e pesquisas 6.4</p> <p>Debate de ideias 6.5</p>	<p>Riscos à integridade acadêmica 7.1</p> <p>Avaliações de aprendizagem injustas 7.2</p> <p>Informação imprecisa 7.3</p> <p>Dependência excessiva da IA 7.5</p>
Grossi et al.[17]	<p>Aprimoramento da prática pedagógica 6.2</p> <p>Acesso à informação 6.7</p>	<p>Riscos à integridade acadêmica 7.1</p> <p>Informação imprecisa 7.3</p>
Continuação na próxima página		

Tabela 9.2 – Continuação

Autor	Benefícios	Riscos
Chung Kwan[7]	<p>Criação de avaliações de aprendizagem 6.1</p> <p>Aprimoramento da prática pedagógica 6.2</p> <p>Oferecimento de tutoria personalizada 6.3</p> <p>Criação de esboços de artigos e pesquisas 6.4</p>	<p>Riscos à integridade acadêmica 7.1</p> <p>Avaliações de aprendizagem injustas 7.2</p> <p>Informação imprecisa 7.3</p>
Continuação na próxima página		

Tabela 9.2 – Continuação

Autor	Benefícios	Riscos
Mai et al.[11]	Criação de avaliações de aprendizagem 6.1	
	Aprimoramento da prática pedagógica 6.2	Riscos à integridade acadêmica 7.1
	Oferecimento de tutoria personalizada 6.3	Avaliações de aprendizagem injustas 7.2
	Criação de esboços de artigos e pesquisas 6.4	Informação imprecisa 7.3
	Debate de ideias 6.5	Dependência excessiva da IA 7.5
	Conscientização ética 6.6	Perpetuação de preconceitos 7.4
	Acesso à informação 6.7	
Continuação na próxima página		

Tabela 9.2 – Continuação

Autor	Benefícios	Riscos
Javaid et al. [14]	<p>Criação de avaliações de aprendizagem 6.1</p> <p>Aprimoramento da prática pedagógica 6.2</p> <p>Oferecimento de tutoria personalizada 6.3</p> <p>Criação de esboços de artigos e pesquisas 6.4</p> <p>Debate de ideias 6.5</p> <p>Acesso à informação 6.7</p>	<p>Riscos à integridade acadêmica 7.1</p> <p>Avaliações de aprendizagem injustas 7.2</p> <p>Informação imprecisa 7.3</p> <p>Perpetuação de preconceitos 7.4</p> <p>Riscos à privacidade 7.6</p>
Hasanein e Sobaih [15]	<p>Criação de avaliações de aprendizagem 6.1</p> <p>Redução da ansiedade</p> <p>Criação de esboços de artigos e pesquisas 6.4</p> <p>Acesso à informação 6.7</p>	<p>Riscos À integridade acadêmica</p> <p>Informação imprecisa 7.3</p> <p>Perpetuação de preconceitos 7.4</p> <p>Dependência excessiva da IA 7.5</p>
Continuação na próxima página		

Tabela 9.2 – Continuação

Autor	Benefícios	Riscos
Ipek et al. [16]	<p>Criação de avaliações de aprendizagem 6.1</p> <p>Oferecimento de tutoria personalizada 6.3</p>	<p>Dependência excessiva da IA 7.5</p> <p>Riscos à integridade acadêmica 7.1</p> <p>Avaliações de aprendizagem injustas 7.2</p> <p>Informação imprecisa 7.3</p> <p>Perpetuação de preconceitos 7.4</p>
Kloeckner et al. [18]	<p>Criação de avaliações de aprendizagem 6.1</p> <p>Aprimoramento da prática pedagógica 6.2</p> <p>Criação de esboços de artigos e pesquisas 6.4</p> <p>Oferecimento de tutoria personalizada 6.3</p>	<p>Riscos À integridade acadêmica 7.1</p> <p>Informação imprecisa 7.3</p> <p>Riscos à privacidade 7.6</p> <p>Perpetuação de preconceitos 7.4</p>

Além disso, o *ChatGPT* facilita a geração de ideias e a organização de pensamentos, o que pode estimular a criatividade e o pensamento crítico dos alunos. Destaca-se em [14] que o uso da inteligência artificial no ambiente educacional permite que os alunos

experimentem diferentes abordagens para a solução de problemas, o que pode enriquecer o processo de aprendizagem. Essa ferramenta oferece um ambiente seguro para a prática e o erro, essencial para o desenvolvimento do pensamento crítico e para a construção de conhecimento.

De acordo com [71], a democratização do acesso ao conhecimento é outro ponto forte do *ChatGPT*. Em muitas regiões, o acesso a recursos educacionais de alta qualidade é limitado, seja por questões econômicas ou geográficas. O *ChatGPT*, ao estar disponível em qualquer dispositivo com acesso à internet, oferece aos estudantes a oportunidade de aprender de maneira mais equitativa, independentemente de sua localização ou condições socioeconômicas. Como apontado por [69], a possibilidade de acessar informações e conteúdos educacionais a qualquer momento e em qualquer lugar amplia as oportunidades de aprendizagem, ou seja promovendo a equidade digital, para aqueles que teriam recursos limitados.

Outro benefício significativo é a potencial redução da carga de trabalho dos professores. Ao automatizar tarefas rotineiras, como a resposta a perguntas frequentes e a correção preliminar de exercícios, o *ChatGPT* permite que os docentes se concentrem em aspectos mais complexos e interativos do ensino. Isso pode resultar em um ambiente de aprendizagem mais dinâmico e atraente, onde os professores têm mais tempo para se dedicar à orientação individualizada e ao desenvolvimento de atividades que incentivem a participação ativa dos estudantes. Como observado em [14], a utilização do *ChatGPT* pode liberar os professores para se concentrarem em atividades de ensino mais criativas e em estratégias pedagógicas inovadoras.

Apesar dos benefícios destacados, é fundamental analisar e debater sobre os riscos associados ao uso do *ChatGPT*. Uma das principais preocupações é a integridade acadêmica, especialmente no que diz respeito ao plágio e à desonestidade acadêmica. Como discutido em [13], o *ChatGPT* pode ser utilizado indevidamente por estudantes para gerar respostas ou textos que são apresentados como sendo de sua própria autoria, sem a devida citação das fontes. Para evitar esse problema, é essencial que as instituições implementem políticas rigorosas de controle de plágio, além da identificação de textos gerados por inteligência artificial e que promovam a conscientização sobre a importância da originalidade e da ética acadêmica. A tabela 9.3 apresenta um resumo das recomendações feitas pelos autores estudados.

A dependência excessiva do *ChatGPT* é outro risco que merece atenção. A ferramenta não deve substituir a interação humana e a troca de ideias entre estudantes e professores, que são fundamentais para o desenvolvimento de habilidades críticas e criativas. Para mitigar esse risco, é necessário promover o uso do *ChatGPT* como um complemento ao ensino tradicional, e não como um substituto. Segundo [18], o uso equilibrado dessa

tecnologia pode evitar que os estudantes se tornem excessivamente dependentes dela, preservando a autenticidade de seu processo de aprendizagem.

A disseminação de informações imprecisas é um ponto a ser considerado, uma vez que o *ChatGPT*, como qualquer outra ferramenta de inteligência artificial, pode produzir respostas baseadas em dados incorretos ou incompletos. Conforme [60], é importante treinar estudantes e professores para que utilizem o *ChatGPT* de maneira crítica, verificando a precisão das informações geradas e desenvolvendo habilidades para avaliar a confiabilidade das fontes.

O uso de ferramentas de inteligência artificial implica na coleta e armazenamento de grandes volumes de dados, o que pode gerar preocupações sobre a segurança dessas informações. Para garantir a privacidade dos estudantes, as instituições de ensino devem adotar políticas claras e transparentes sobre o uso de dados, conforme sugerido por [1]. Além disso, é fundamental que esses tipos de plataformas estejam em conformidade com as regulamentações de proteção de dados, garantindo assim um ambiente seguro para o uso dessas tecnologias.

Em resumo, o *ChatGPT* apresenta uma série de benefícios significativos para o ambiente educacional, como a facilitação do acesso ao conhecimento, o apoio ao desenvolvimento da criatividade e do pensamento crítico dos estudantes, e a democratização das oportunidades de aprendizagem. Entretanto, é imprescindível que as instituições de ensino adotem medidas cabíveis, garantindo o uso responsável e ético dessa tecnologia. A implementação de diretrizes claras e a promoção de práticas educativas inovadoras são passos fundamentais para que o *ChatGPT* se torne uma ferramenta eficaz no aprimoramento do processo ensino-aprendizagem.

Apesar dos resultados positivos apresentados, é necessário aprofundar a análise dos impactos a longo prazo e explorar as melhores estratégias para integrar essa tecnologia no cotidiano acadêmico de forma coesa. O próximo capítulo abordará essas conclusões, discutindo como os benefícios e riscos identificados podem influenciar a prática pedagógica e a formação dos estudantes.

Tabela 9.3: Tabela contendo o resumo das recomendações feitas pelos autores para os riscos apontados pelas publicações estudadas.

Riscos	Recomendações
Riscos a integridade acadêmica 7.1	<p>Conferência de fontes 7.1</p> <p>Ferramentas identificadoras de plágio baseadas em inteligência artificial 7.1</p> <p>Regras claras 7.1</p> <p>Capacitação de docentes e discentes 7.1</p> <p>Sistemas de monitoramento 7.1</p>
Avaliações de aprendizagem injustas 7.2	<p>Avaliações formativas 7.2</p> <p>Promoção da análise e criatividade 7.2</p> <p>Provas totalmente livres de recursos digitais 7.2</p> <p>Incorporação de recursos midiáticos 7.2</p>
Informação imprecisa 7.3	<p>Capacitação de docentes e discentes 7.3</p> <p>Conferência de fontes 7.3</p>
Perpetuação de preconceitos 7.4	<p>Feedback humano no treinamento da IA 7.4</p> <p>Diversidade nos dados de treinamento 7.4</p> <p>Sistemas de monitoramento 7.4</p>
Dependência excessiva da IA 7.5	<p>Sistemas de monitoramento 7.5</p> <p>Promoção da análise e criatividade 7.5</p> <p>Capacitação de docentes e discentes 7.5</p>
Riscos à privacidade 7.6	<p>Anonimização dos dados 7.6</p> <p>Capacitação de docentes e discentes 7.6</p> <p>Regras claras 7.6</p> <p>Sistemas de monitoramento 7.6</p>

Capítulo 10

Conclusões

Ao final dessa pesquisa, foi possível concluir que as inteligências artificiais, especialmente o *ChatGPT*, estão mudando profundamente a forma de estudar e ensinar. As investigações permitiram levantar diversos benefícios e riscos associados ao seu uso, bem como recomendações que permitiriam reduzir os riscos observados, otimizando o uso dessas ferramentas por parte dos docentes e discentes.

Como apresentado no capítulo 6, são vários os benefícios do uso de ferramentas de Inteligência artificial no contexto acadêmico. Criação de avaliações de aprendizagem 6.1, aprimoramento da prática pedagógica 6.2, oferecimento de tutoria personalizada 6.3, criação de esboços de artigos e pesquisas 6.4, debate de ideias 6.5, conscientização ética 6.6, acesso à informação 6.7 e redução da ansiedade 6.8 são apenas alguns dos benefícios observados, mesmo com poucos anos de uso do *ChatGPT*.

No entanto, é importante ressaltar que o uso do *ChatGPT* no ambiente acadêmico também apresenta algumas limitações. Há riscos à integridade acadêmica 7.1 e à privacidade de docentes e discentes 7.6, bem como problemas relacionados a avaliações 7.2, preconceitos 7.4, dependência 7.5 e informações imprecisas 7.3.

A ética relacionada ao uso de ferramentas baseadas em inteligência artificial também precisa ser observada com cuidado. De acordo com [14] há receios sobre a substituição dos professores por sistemas automatizados, embora, o próprio [14] demonstre como essa substituição não ocorrerá. Torna-se, portanto, de acordo com [62], fundamental refletir sobre os impactos sociais e educacionais dessas ferramentas, buscando equilibrar o uso da inteligência artificial com a preservação do papel do professor como mediador do conhecimento, uma das recomendações apontadas no capítulo 8.1.

Contudo, é de entendimento de [13] que as instituições de ensino deveriam, antes de proibir o uso de tais ferramentas, ensinar como aproveitar essa ferramenta de inteligência artificial para ajudar nos estudos e pesquisas de maneira responsável, sem abdicar da ética e das regras acadêmicas, outra recomendação apresentada no capítulo 8.3.

Além das recomendações ditas acima, foram abordados outras recomendações propostas pelos autores para mitigar os riscos apresentados, como conscientização, capacitação dos usuários e anonimização de dados para proteção da privacidade (8.6); diversificação de dados e feedback humano no treinamento para barrar a perpetuação de preconceitos (8.5); sugerir atividades que promovam o pensamento crítico, de modo a evitar uma dependência excessiva de IA (8.4); e adotar avaliações formativas e questões que estimulem análise e criatividade evitando avaliações injustas (8.2).

No capítulo 9 são sintetizados, em forma de tabelas, todos os riscos (7) e as recomendações (8) propostas por cada autor, proporcionando uma visão estruturada do tema no contexto acadêmico. Além disso é abordado como o *ChatGPT* no ambiente acadêmico tem promovido uma transformação na forma como o conhecimento é construído. A ferramenta se destaca por facilitar a compreensão de conceitos complexos, democratizar o acesso ao aprendizado e reduzir a carga de trabalho dos professores, permitindo um ensino mais personalizado e dinâmico. Foi observado que essas análises podem ser aplicadas para outras ferramentas de inteligência artificial semelhantes.

Afinal, as perspectivas futuras do uso de ferramentas baseadas em inteligência artificial no ambiente acadêmico são, segundo [59], promissoras. Espera-se o desenvolvimento de sistemas ainda mais avançados e personalizados, capazes de auxiliar no ensino e aprendizagem de forma ainda mais eficaz. Essas ferramentas podem se tornar cada vez mais integradas ao ambiente acadêmico, contribuindo para a melhoria da qualidade do ensino e para a formação de estudantes mais preparados para os desafios do mercado de trabalho.

Segundo [18], o avanço da tecnologia torna possível vislumbrar a personalização do ensino por meio do uso de sistemas inteligentes que se adaptam às necessidades individuais dos estudantes. Espera-se, ainda segundo [18], que o desenvolvimento de sistemas mais avançados sejam capazes de compreender e responder a perguntas complexas com maior precisão. De acordo com [69], esses sistemas poderão identificar, não apenas erros gramaticais, mas também falhas lógicas e argumentativas.

É por isso que, segundo [3], ao olhar para o futuro, é fundamental que haja um debate contínuo sobre os impactos e desafios das ferramentas baseadas em inteligência artificial no ambiente acadêmico. É necessário adaptar constantemente as práticas educacionais às novas tecnologias disponíveis, considerando sempre os aspectos éticos e sociais envolvidos. A evolução dessas ferramentas deve ser acompanhada de perto, para que sejam aproveitadas da melhor forma possível, visando sempre o aprimoramento do ensino e da pesquisa acadêmica.

De acordo com [67], a integração com outras tecnologias emergentes, como realidade virtual e aumentada, pode potencializar ainda mais os benefícios dessas ferramentas no processo de construção do conhecimento. A combinação dessas tecnologias pode propor-

cionar experiências de aprendizagem imersivas e interativas, ampliando as possibilidades de ensino e tornando-o ainda mais atrativo para os alunos.

Essa breve análise do uso do *ChatGPT* no ambiente acadêmico permitiu elencar diversos pontos positivos (benefícios 6) e diversos pontos negativos (riscos 7). É nesse sentido que o artigo [14] determina quatro paradoxos envolvendo o uso do *ChatGPT*: o *ChatGPT* é "amigo", mas "inimigo"; "capaz", mas "dependente"; "acessível", mas "restritivo"; "popular" mesmo "banido".

Por mais interessantes que sejam esses pontos positivos, um risco relevante, ainda pouco explorado pelos artigos analisados, é a possibilidade de manipulação do conhecimento gerado por essas tecnologias de inteligência artificial de acordo com os interesses de seus desenvolvedores e proprietários. Como os modelos de IA são treinados com base em dados selecionados e ajustados por empresas e instituições com diferentes objetivos, há o risco de viés na informação disponibilizada, influenciando percepções, decisões acadêmicas e até mesmo políticas educacionais. Essa questão reforça a necessidade de maior transparência nos processos de treinamento dessas tecnologias, além da necessidade do desenvolvimento de mecanismos para garantir a pluralidade e a imparcialidade do conhecimento oferecido por esses sistemas.

Por fim, é importante lembrar que o *ChatGPT* é uma ferramenta. Ele não é dotado de consciência, tampouco de vontade própria. Ele apenas calcula a sequência de palavras mais prováveis de aparecer dado um determinado contexto. Cabe à humanidade escolher os usos que serão dados a essa ferramenta. E as universidades, como fomentadoras e impulsionadoras do conhecimento, têm o dever de orientar a sociedade sobre os benefícios e riscos que essas tecnologias apresentam, bem como destinar os esforços para promover o melhor aproveitamento dessas ferramentas no processo de ensino-aprendizagem, seja através do ensino, da pesquisa ou da extensão.

Capítulo 11

Trabalhos Futuros

Por ser uma tecnologia recente, sendo disponibilizada a partir de 2022 de acordo com [6], não existem dados da eficácia do uso dessa ferramenta no Brasil. Recomenda-se para trabalhos futuros fazer uma análise comparativa entre instituições de ensino que usam essa ferramenta e as que não usam, para verificar o desempenho dessa ferramenta nesses ambientes, obtendo assim resultados mais acertos acerca dos benefícios e dos riscos da sua utilização.

Outra pesquisa futura interessante, seria a análise dos feedbacks personalizados que o *ChatGPT* gera para alunos e professores, analisando sua eficácia e originalidade para verificar os potenciais dessa ferramenta no ambiente acadêmico. Ainda em cima dos feedbacks poderia ser analisado vieses e tendências para verificar se o algoritmo está perpetuando preconceitos ou enviesando autores para um determinado caminho.

Uma linha de pesquisa interessante é o uso do *ChatGPT* e a análise crítica dos dados gerados, trazendo dados de pessoas que potencializaram a análise crítica e o melhor desempenho na produção de matérias para a academia, ao mesmo tempo trazendo possíveis riscos de dependência do uso dessa ferramenta. Caso aconteça essa dependência, também começa a analisar se isso é algo ruim ou bom para o ambiente acadêmico, pois assim como somos dependentes da *Internet* para comunicação e do Google para fazer buscas de artigos e trabalhos para produzir material acadêmico, o *ChatGPT* pode vir a ser uma nova tecnologia que ajudará a desenvolver o potencial para esses trabalhos.

Outro ramo que poderia surgir é engenharia de *prompt*, como matéria obrigatória em cursos de graduação, se o *ChatGPT*, ou alguma ferramenta de inteligência artificial for amplamente usada pelas instituições de ensino. Assim como é ensinado nas faculdades a maneira de ir atrás de boas ferramentas de pesquisa, sites, artigos e usar ferramentas para potencializar a produção de material acadêmico, seria importante investigar e estudar como ensinar diferentes áreas a usar bons comandos para ter a ferramenta de inteligência artificial trabalhando da melhor e mais eficaz forma possível. Além de entender qual

ferramenta de inteligência artificial é mais adequada para resolver o problema em questão.

Um ramo interessante de pesquisa é o aprofundamento e a análise ética do uso de ferramentas baseadas em inteligência artificial no processo de ensino-aprendizagem. Isso inclui investigar os impactos a longo prazo na autonomia dos estudantes, na equidade, no acesso à educação e na construção do pensamento crítico. Deve-se, também, explorar diretrizes e regulamentações que garantam o uso responsável dessas tecnologias, evitando o desenvolvimento de dependências e preservando o papel ativo dos educadores na construção do conhecimento dos alunos.

Outra possibilidade para trabalhos futuros é a análise da eficácia da personalização do processo de ensino-aprendizagem para diferentes perfis de estudantes, incluindo pessoas com deficiência, autismo e outras limitações. Investigar se as ferramentas de inteligência artificial conseguem atender adequadamente às necessidades específicas desses grupos, promovendo a inclusão e a equidade na educação, além de avaliar as possíveis barreiras tecnológicas e metodológicas que possam limitar a adaptação dessas soluções a contextos diversos, garantindo que a personalização não apenas potencialize o processo de ensino-aprendizagem, mas também respeite as particularidades de cada indivíduo.

Referências

- [1] Lyra, M. W. A.: *Revolução Industrial 4.0 na educação: explorando o impacto transformador das Inteligências Artificiais no ensino—Uma abordagem introdutória às novas ferramentas*. <https://app.uff.br/riuff/handle/1/30383>. 1, 31, 49, 62
- [2] Giraffa, L. e P. Khol-Santos: *Inteligência Artificial e Educação: conceitos, aplicações e implicações no fazer docente*. *Educação em Análise*, 2023, 2023. <https://ojs.uel.br/revistas/uel/index.php/educanalise/article/view/48127>. 1, 32, 47
- [3] Bicalho, G. C.: *A Presença da Inteligência Artificial na Educação: as vantagens e divergências do uso do CHATGPT no meio acadêmico*. Em *Anais do Congresso Nacional Universidade, EAD e ...*, 2023. 1, 42, 43, 44, 49, 65
- [4] Pereira, I. S. D. e S. A. de Moura: *Explorações teóricas e oportunidades de integração curricular do letramento em inteligência artificial (IA) na educação*. Preprints SciELO, 7294, 2021. <https://preprints.scielo.org/index.php/scielo/preprint/view/7294>. 1, 2, 3, 31, 39, 47, 48
- [5] Abdullah, Malak, Alia Madain e Yaser Jararweh: *ChatGPT: Fundamentals, Applications and Social Impacts*. novembro 2022. <https://ieeexplore.ieee.org/abstract/document/10062688>. 1, 11, 17, 21
- [6] Wu, Tianyu, Shizhu He, Jinping Liu, Siqi Sun, Kang Liu, Qing Long Han e Yang Tang: *A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development*. página 15, março 2023. <https://ieeexplore-ieee-org.ez54.periodicos.capes.gov.br/document/10113601>. 1, 3, 9, 20, 67
- [7] Lo, Chung Kwan: *What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature*. *Education Sciences*, 13(4), 2023, ISSN 2227-7102. <https://www.mdpi.com/2227-7102/13/4/410>. 2, 5, 50, 51, 57
- [8] Ferreira, R. C. V. e G. H. M. Garcia: *O surgimento do Chat GPT e a insegurança sobre o futuro dos trabalhos acadêmicos*. *Cadernos de Direito Actual*, 2023. <https://www.cadernosdedereitoactual.es/ojs/index.php/cadernos/article/view/917>. 2, 30, 37, 39
- [9] Cavalcante, I. F.: *Reflexões sobre a produção do conhecimento em face da Inteligência Artificial*. *Revista de Educação*, 2023. <https://puccampinas.emnuvens.com.br/reveducacao/article/view/8671>. 2, 40, 49

- [10] Silva, KR da e LS de Oliveira Barbosa: *INTELIGÊNCIA ARTIFICIAL E SEUS IMPACTOS NA EDUCAÇÃO: UMA REVISÃO SISTEMÁTICA*. Recima21, 1(1):1–10, 2023, ISSN 2675-6218. <https://recima21.com.br/index.php/recima21/article/view/4353>. 2, 30, 32, 38, 41, 43, 48
- [11] Mai, Duong Thi Thuy, Can Van Da e Nguyen Van Hanh: *The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach*. Frontiers in Education, 9, 2024, ISSN 2504-284X. <https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2024.1328769>. 2, 5, 34, 37, 38, 43, 46, 50, 52, 58
- [12] OpenAI: *GPT-4 Technical Report*. março 2023. <https://arxiv.org/abs/2303.08774>. xi, 3, 26, 27, 28, 29
- [13] Sok, Sarin e Kimkong Heng: *ChatGPT for Education and Research: A Review of Benefits and Risks*, março 2023. <https://papers.ssrn.com/abstract=4378735>, acesso em 2024-07-24. 5, 34, 36, 37, 38, 40, 41, 42, 45, 46, 47, 48, 51, 56, 61, 64
- [14] Javaid, Mohd, Abid Haleem, Ravi Pratap Singh, Shahbaz Khan e Ibrahim Haleem Khan: *Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system*. BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 3(2):100115, 2023, ISSN 2772-4859. <https://www.sciencedirect.com/science/article/pii/S2772485923000327>. 5, 37, 44, 45, 46, 47, 49, 52, 53, 54, 59, 60, 61, 64, 66
- [15] Hasanein, Ahmed M. e Abu Elnasr E. Sobaih: *Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives*. European Journal of Investigation in Health, Psychology and Education, 13(11):2599–2614, novembro 2023, ISSN 2174-8144. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10670526/>, acesso em 2024-08-12. 5, 37, 44, 51, 52, 59
- [16] İpek, Ziyaeddin Halid, Ali İbrahim Can Gözümlü, Stamatios Papadakis e Michail Kallogiannakis: *Educational Applications of the ChatGPT AI System: A Systematic Review Research*. Educational Process: International Journal, junho 2023. <https://www.edupij.com/index/arsiv/59/305/educational-applications-of-the-chatgpt-ai-system-a-systematic-review-research>, acesso em 2024-08-13. 5, 45, 52, 60
- [17] Grossi, M. G. R., R. V. Rosa e C. de Aguiar: *Contribuições da inteligência artificial para a educação: uma entrevista com o ChatGPT*. Synthesis: Revista Digital FAPAM, 10(1):1–10, 2023. <https://periodicos.fapam.edu.br/index.php/synthesis/article/view/635>. 5, 38, 42, 49, 56
- [18] Kloeckner, F. L. e E. M. Giordani: *Inteligência Artificial nos processos de ensino-aprendizagem no ensino superior: uma revisão narrativa*. Revista Contribuciones a las Ciencias, 2023. <https://ojs.revistacontribuciones.com/ojs/index.php/clcs/article/view/1736>. 5, 41, 44, 45, 47, 48, 60, 61, 65
- [19] Wei, Qiuhong: *Evaluation of ChatGPT's Performance in Providing Treatment Recommendations for Pediatric Diseases*. Pediatric Discovery, 2023. 7

- [20] Mitsuyama, Yasuhito: *Comparative Analysis of ChatGPT's Diagnostic Performance With Radiologists Using Real-World Radiology Reports of Brain Tumors*. 2023. 7
- [21] Al-Obaydi, Liqaa Habeb: *ChatGPT and the General Concepts of Education: Can Artificial Intelligence-Driven Chatbots Support the Process of Language Learning?* International Journal of Emerging Technologies in Learning (Ijet), 2023. 7
- [22] Frieder, Simon, Luca Pinchetti, Alexis Chevalier, Ryan Rhys Griffiths, Salvatori Tommaso, Thomas Lukasiewicz, Philipp Petersen e Julius Berner: *Mathematical Capabilities of ChatGPT*. página 37, julho 2023. <https://arxiv.org/abs/2301.13867>. x, 7, 26, 27
- [23] Sütçüoğlu, Bengü Mutlu: *Appropriateness of Premature Ovarian Insufficiency Recommendations Provided by ChatGPT*. Menopause the Journal of the North American Menopause Society, 2023. 8
- [24] Golden, Willie: *ChatGPT: A Trusted Source?* Irish Journal of Technology Enhanced Learning, 2023. 8
- [25] Aggarwal, Karan, Maad M. Mijwil, Sonia, Abdel Hameed Al-Mistarehi, Safwan Alomari, Murat Gök, Anas M. Zein Alaabdin e Safaa H. Abdulrhman: *Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning*. Iraqi Journal For Computer Science and Mathematics, 3(1):115–123, janeiro 2022. <https://journal.esj.edu.iq/index.php/IJCM/article/view/1009>, 10
- [26] Goodfellow, Ian, Yoshua Bengio e Aaron Courville: *Deep Learning*. MIT Press, 2016. 9, 10, 11, 13, 14, 15, 16
- [27] Kersting, Kristian: *Machine Learning and Artificial Intelligence: Two Fellow Travelers on the Quest for Intelligent Behavior in Machines*. Sec. Machine Learning and Artificial Intelligence, 1, novembro 2018. <https://doi.org/10.3389/fdata.2018.00006>. 9, 10
- [28] TURING, A. M.: *I.—COMPUTING MACHINERY AND INTELLIGENCE*. Mind, LIX(236):433–460, outubro 1950, ISSN 0026-4423. <https://doi.org/10.1093/mind/LIX.236.433>, _eprint: <https://academic.oup.com/mind/article-pdf/LIX/236/433/30123314/lix-236-433.pdf>. 10
- [29] Biever, Celeste: *ChatGPT broke the Turing test - the race is on for new ways to assess AI*. Nature (London), 619(7971):686, 2023, ISSN 1476-4687. <https://www-nature.ez54.periodicos.capes.gov.br/articles/d41586-023-02361-7>, Place: England. 10
- [30] Abbass, Hussein: *Editorial: What is Artificial Intelligence?* páginas 94–95, 2021. <https://ieeexplore-ieee-org.ez54.periodicos.capes.gov.br/document/9523786>. 10
- [31] Mohri, Mehryar, Afshin Rostamizadeh e Ameet Talwalkar: *Foundations of Machine Learning*. The MIT Press, 2012, ISBN 0-262-01825-X. ix, 10, 11, 12, 17

- [32] Wagh, Sameer, Shruti Tople, Fabrice Benhamouda, Eyal Kushilevitz, Prateek Mittal e Tal Rabin: *Falcon: Honest-Majority Maliciously Secure Framework for Private Deep Learning*. Proceedings on Privacy Enhancing Technologies, 2020. 11
- [33] Dash, S., S. Nayak e D. Mishra: *A review on machine learning algorithms*. páginas 495–507, 2020. 12
- [34] Grau, M., R. Centelles e F. Freitag: *On-device training of machine learning models on microcontrollers with a look at federated learning*. 2021. 12
- [35] Parthasarathy, Sudhaman e S.T. Padmapriya: *Understanding algorithm bias in artificial intelligence-enabled ERP software customization*. Journal of Ethics in Entrepreneurship and Technology, 3(2):79–93, janeiro 2023, ISSN 2633-7436, 2633-7444. <https://doi.org/10.1108/JEET-04-2023-0006>, acesso em 2024-09-12, Publisher: Emerald Publishing Limited. 13, 14
- [36] Wang, Xukang, Ying Cheng Wu, Xueliang Ji e Hongpeng Fu: *Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices*. Frontiers in Artificial Intelligence, 7, maio 2024, ISSN 2624-8212. <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1320277/full>, acesso em 2024-09-12, Publisher: Frontiers. 14
- [37] Fazil, Abdul Wajid, Musawer Hakimi e Amir Krór Shahidzay: *A COMPREHENSIVE REVIEW OF BIAS IN AI ALGORITHMS*. Nusantara Hasana Journal, 3(8):1–11, 2023, ISSN 2798-1428. <https://nusantarahasanajournal.com/index.php/nhj/article/view/1052>, acesso em 2024-09-12, Number: 8. 14
- [38] LeCun, Yann, Yoshua Bengio e Geoffrey E. Hinton: *Deep Learning*. Nature, 2015. 16
- [39] Levy, Joshua, Christopher M. Navas, Chandra J, Brock C. Christensen, Louis Vaickus, Michael A. Curley, William D. Chey, Jason Baker e Eric D. Shah: *Video-Based Deep Learning to Detect Dyssynergic Defecation With 3D High-Definition Anorectal Manometry*. 2021. 16
- [40] Wang, Haoriqin, Mingyang Jiang, Jianhong Qi, Xinhong Zhang, Qinghua Wang, Yuxin Zhou, Mingyu Bai, Lisha Li e Zhili Pei: *Application of Deep Learning in Text Mining*. 2014. 16
- [41] Saxe, Andrew, Stephanie Nelli e Christopher Summerfield: *If deep learning is the answer, what is the question?* Nature Reviews Neuroscience, 22(1):55–67, 2021. Publisher: Nature Publishing Group UK London. 16
- [42] Zhao, Wayne, kun Zhou, Juniy Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian Yun Nie e Ji Rong Wen: *A Survey of Large Language Models*. página 97, setembro 2023. <https://arxiv.org/abs/2303.18223>. 17, 19

- [43] Kallens, Pablo Contreras, Ross Deans Kristensen-McLachlan e Morten H. Christiansen: *Large Language Models Demonstrate the Potential of Statistical Learning in Language*. Cognitive Science, 2023. 17
- [44] Naveed, Humza, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes e Ajmal Mian: *A Comprehensive Overview of Large Language Models*, 2024. _eprint: 2307.06435. 17, 18
- [45] Brill, Eric e Raymond J. Mooney: *An Overview of Empirical Natural Language Processing*. AI Magazine, 18(4):13, dezembro 1997. <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1318>. 17
- [46] Nadkarni, Prakash M, Lucila Ohno-Machado e Wendy W Chapman: *Natural language processing: an introduction*. Journal of the American Medical Informatics Association, 18(5):544–551, setembro 2011, ISSN 1067-5027. <https://doi.org/10.1136/amiajnl-2011-000464>, _eprint: <https://academic.oup.com/jamia/article-pdf/18/5/544/5962687/18-5-544.pdf>. 17, 21
- [47] Allen, James F.: *Natural Language Processing*. Em *Encyclopedia of Computer Science*, páginas 1218–1222. John Wiley and Sons Ltd., GBR, 2003, ISBN 0-470-86412-5. <https://dl.acm.org/doi/abs/10.5555/1074100.1074630>. 17, 18
- [48] AYDIN, Ömer e Enis KARAARSLAN: *Is ChatGPT Leading Generative AI? What is Beyond Expectations?* Academic Platform Journal of Engineering and Smart Systems, 11(3):118–134, 2023. Publisher: Akademik Perspektif Derneği. 18
- [49] Jaiswal, Ashish, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Debapriya Banerjee e Fillia Makedon: *A Survey on Contrastive Self-Supervised Learning*. Technologies, 9(1), 2021, ISSN 2227-7080. <https://www.mdpi.com/2227-7080/9/1/2>. 18
- [50] Yenduri, Gokul, Ramalingam Murugan, Chemmalar Govardanan, Supriya Y, Gautam Srivastava, Praveen Reddy, Deepti Raj, Rutvij Jhaveri, Prabadevi B, Weizheng Wang, Athanasios Vasilakos e Thippa Godekallu: *Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions*. maio 2023. https://www.researchgate.net/publication/370869544_Generative_Pre-trained_Transformer_A_Comprehensive_Review_on_Enabling_Technologies_Potential_Applications_Emerging_Challenges_and_Future_Directions. 19
- [51] Han, Xu, Zhengyan Zhang, Ning Ding, Yuxian Gu, Xiao Liu, Yuqi Huo, Jiezhong Qiu, Yuan Yao, Ao Zhang, Liang Zhang, Wentao Han, Minlie Huang, Qin Jin, Yanyan Lan, Yang Liu, Zhiyuan Liu, Zhiwu Lu, Xipeng Qiu, Ruihua Song, Jie Tang, Ji Rong Wen, Jinhui Yuan, Wayne Xin Zhao e Jun Zhu: *Pre-trained models: Past, present and future*. AI Open, 2:225–250, 2021, ISSN 2666-6510. <https://www.sciencedirect.com/science/article/pii/S2666651021000231>. 21

- [52] Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever e Dario Amodei: *Language Models are Few-Shot Learners*. Relatório Técnico, 2020. <https://arxiv.org/abs/2005.14165>, [_eprint: 2005.14165](https://arxiv.org/abs/2005.14165). 21
- [53] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser e Illia Polosukhin: *Attention is All you Need*. Em Guyon, I., U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan e R. Garnett (editores): *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf. ix, 21, 22, 23, 24, 25
- [54] Grefenstette, Gregory: *Tokenization*. Em Halteren, Hans van (editor): *Syntactic Wordclass Tagging*, páginas 117–133. Springer Netherlands, Dordrecht, 1999, ISBN 978-94-015-9273-4. https://doi.org/10.1007/978-94-015-9273-4_9. 24
- [55] Kowsari, Kamran, Kiana Jafari Meimandi, Mojtaba Heidarysafa, Sanjana Mendu, Laura Barnes e Donald Brown: *Text Classification Algorithms: A Survey*. Information, 10(4), 2019, ISSN 2078-2489. <https://www.mdpi.com/2078-2489/10/4/150>. 24, 25
- [56] *OpenAI Tokenizer*. <https://platform.openai.com/tokenizer>, acesso em 2023-12-04. 25
- [57] Liu, Zhijie, Yutian Tang, Xiapu Luo, Yuming Zhou e Liang Zhang: *No Need to Lift a Finger Anymore? Assessing the Quality of Code Generation by ChatGPT*. página 29, agosto 2023. <https://arxiv.org/abs/2308.04838>. 28
- [58] Figueiredo, L. de O. e A. M. Z. Lopes: *Desafios e impactos do uso da Inteligência Artificial na educação*. Educação Online, 2023. <http://educacaoonline.edu.puc-rio.br/index.php/eduonline/article/view/1506>. 31, 39, 40
- [59] Nascimento, C. C.: *Inteligência artificial no ensino superior: da transformação digital aos desafios da contemporaneidade*. <https://www.editorapublicar.com.br/ojs/index.php/publicacoes/article/view/788>. 31, 32, 38, 41, 42, 44, 48, 54, 65
- [60] Vieira, A. L. F., M. C. Z. Amorim e E. Cunha: *Proposta de avaliação da percepção dos impactos da inteligência artificial generativa na educação superior*. Em *Anais do XIV Simpósio...*, 2023. <https://sol.sbc.org.br/index.php/stil/article/view/25480>. 31, 40, 43, 51, 62
- [61] Alcoforado, J. G. D.: *Uso do ChatGPT na pesquisa acadêmica*. <https://bdta.ufra.edu.br/jspui/handle/123456789/3345>. 32, 40, 42, 48

- [62] Brasão, M. dos Reis: *Perspectivas da Inteligência Artificial na Educação: Modelagem computacional, semântica e ChatGPT*. 2023. <https://editorapantanal.com.br/ebooks/2023/educacao-dilemas-contemporaneos-volume-xvii/ebook.pdf#page=19>. 32, 36, 37, 64
- [63] Jacinto, E., A. L. C. Bertoncini e M. C. F. D. C. Ames: *Os desafios do uso do ChatGPT no ensino e pesquisa em Administração: uma discussão baseada na ética das virtudes*. https://www.researchgate.net/profile/Ana-Correa-Bertoncini/publication/374951164_Os_desafios_do_uso_do_ChatGPT_no_ensino_e_pesquisa_em_Administracao_uma_discussao_baseada_na_etica_das_virtudes/links/65392aaf73a2865c7ad30b26/Os-desafios-do-uso-do-ChatGPT-no-ensino-e-pesquisa-em-Administracao-uma-discussao.pdf. 33, 39, 42
- [64] Olivier, Bert: *Artificial Intelligence (AI) and being human: What is the difference?* Acta Academica: Critical views on society, culture and politics, 49(1):2–21, 2017. 33
- [65] Raja, Rahat e PC Nagasubramani: *Impact of modern technology in education*. Journal of Applied and Advanced Research, 3(1):33–35, 2018. 34
- [66] Williams, Nigel, Stanislav Ivanov e Dimitrios Buhalis: *Algorithmic Ghost in the Research Shell: Large Language Models and Academic Knowledge Creation in Management Research*, 2023. _eprint: 2303.07304. 35
- [67] Maarouf, A. C. R.: *A responsabilidade civil pelo uso do ChatGPT: uma análise dos reflexos jurídicos causados pela utilização da inteligência artificial*. <https://lume.ufrgs.br/handle/10183/259783>. 39, 42, 65
- [68] Yepes, I. e A. Fiorin: *ChatGPT como Assistente de Ensino na Disciplina de Estruturas de Dados*. Em *Anais do Encontro Anual de Tecnologia da...*, 2023. <http://168.228.253.6:8080/index.php/2019/article/view/308>. 43, 49
- [69] Souza, MNM de e PVP de Lima: *DO GPT 3 AO CHATGPT: POTENCIALIDADES E ALERTAS NO ENFOQUE DA PRODUÇÃO ACADÊMICA BRASILEIRA*. Em *Boletim de ...*, 2023. <https://revista.ioles.com.br/boca/index.php/revista/article/view/2732>. 46, 47, 61, 65
- [70] Bennett, Randy e Rbennett@ets Org: *Formative assessment: A critical review*. Assessment in Education: Principles, Policy & Practice, 18, fevereiro 2011. 51
- [71] Dwivedi, Yogesh, Nir Kshetri, Emma Slade, Laurie Hughes, Anand Jeyaraj, Arpan Kar, Abdullah Baabdullah, Alex Koochang, Vishnupriya Raghavan, Manju Ahuja, Hanaa Albana, Mousa Albashrawi, Adil Al-Busaidi, Janarthanan Balakrishnan, Yves Barlette, Sriparna Basu, Indranil Bose, Laurence Brooks, Dimitrios Buhails, Lemuria Carter e Ryan Wright: *So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy*. página 63, agosto 2023. <https://www.sciencedirect.com/science/article/pii/S0268401223000233?via%3Dihub>. 61