PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS
ESCOLA POLITÉCNICA E DE ARTES
GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

DO ELIZA AO CHATGPT: HISTÓRIA E EVOLUÇÃO DA INTELIGÊNCIA ARTIFICIAL

MATHEUS AFONSO BATISTA DA SILVA

GOIÂNIA
2024

MATHEUS AFONSO BATISTA DA SILVA

DO ELIZA AO CHATGPT: HISTÓRIA E EVOLUÇÃO DA INTELIGÊNCIA ARTIFICIAL

Trabalho de Conclusão de Curso apresentado à Escola Politécnica e de Artes, da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para a obtenção do título de Bacharel em ciências da computação.

Orientador(a):

Prof.(a) Me.(a) Lucilia Gomes Ribeiro

Banca examinadora:

Prof. Me. Fernando Gonçalves Abadia

Prof. Dra. Solange da Silva

GOIÂNIA

2024

MATHEUS AFONSO BATISTA DA SILVA

DO ELIZA AO CHATGPT: HISTÓRIA E EVOLUÇÃO DA IA

Trabalho de Conclusão de Curso aprovado em sua forma final pela Escola Politécnica e de Artes, da Pontifícia Universidade Católica de Goiás, para obtenção do título de Bacharel em ciências da computacão, em _____/_____/_____.

_____

Orientador(a): Prof. Ma. Lucilia Gomes Ribeiro

_____

Prof. Me. Fernando Gonçalves Abadia

_____

Prof. Dr. Solange da Silva

GOIÂNIA

2024

Dedico esse projeto de conclusão de curso em primeiro lugar a minha mãe, Rivia Karine da Silva e ao meu pai Clevio Robson da Silva, que esteve ao meu lado em todos os momentos me incentivando e apoiando. Em segundo lugar a minha noiva, Ana Carolina Rodrigues, que constantemente acreditou em mim e me ajudou nessa jornada. Por último a todas as pessoas que amo e que fizeram parte dessa etapa da minha vida.

AGRADECIMENTOS

Aos professores da Pontifícia Universidade Católica de Goiás, principalmente a minha orientadora, Prof. Lucilia Gomes Ribeiro por estar presente durante essa etapa do curso, pela paciência e o auxílio na elaboração deste TCC

“Feliz aquele que transfere o que sabe e aprende o que ensina”.

Cora Coralina

RESUMO

O objetivo geral deste trabalho é analisar como a tecnologia tornou possível treinar e executar modelos como o ChatGPT. Desde a ELIZA, um dos primeiros chatbots criado por Joseph Weizenbaum no MIT entre 1964 e 1966, até os modelos modernos como o GPT-3.5 e GPT-4 desenvolvidos pela OpenAI, o estudo aborda também a importância do processamento de linguagem natural (PLN), redes neurais artificias e convolucionais e o aprendizado profundo e como são tecnologias fundamentais para avanços em áreas como visão computacional e PLN, apesar de enfrentarem desafios como a necessidade de grandes conjuntos de dados e alta demanda computacional. A arquitetura Transformer revolucionou a forma como os modelos de atenção são utilizados, impulsionando a IA no século XXI. A IA pode ser classificada em IA fraca, projetada para tarefas específicas, e IA forte, que busca imitar a inteligência humana. O aprendizado de máquina, um subcampo da IA, se concentra no desenvolvimento de sistemas que podem aprender e melhorar a partir de experiências passadas. O trabalho também destaca a importância de abordar questões éticas e de privacidade, especialmente com o uso de dados sensíveis e a necessidade de promover o uso responsável da tecnologia. A OpenAI, uma empresa líder no campo, está trabalhando para resolver esses problemas e promover a inteligência artificial geral de forma segura e benéfica para todos. Inovações como o ChatGPT estão empurrando as fronteiras do que é possível. Desafiando a compreensão e as capacidades de uma máquina e estabelecendo uma nova fronteira sobre o que IA será capaz no futuro.

Palavras-chaves: Inteligência artificial. Chatgpt. GPT. Tecnologia. Eliza.

ABSTRACT

The overall goal of this paper is to analyze how technology has become possible to train and run models like ChatGPT. From ELIZA, one of the first chatbots created by Joseph Weizenbaum at MIT between 1964 and 1966, to modern models like GPT-3.5 and GPT-4 developed by OpenAI, the study also addresses the importance of natural language processing (NLP), artificial and convolutional neural networks, and deep learning and how they are fundamental technologies for advances in areas like computer vision and NLP, despite facing challenges such as the need for large data sets and high computational demand. The Transformer architecture revolutionized the way attention models are used, propelling AI into the 21st century. AI can compete in weak AI, designed for specific tasks, and strong AI, which seeks to imitate human intelligence. Machine learning, a subfield of AI, focuses on developing systems that can learn and improve from past experiences. The work also highlights the importance of addressing ethical and privacy concerns, especially when dealing with sensitive data, and the need to promote responsible use of technology. OpenAI, a leading company in the field, is working to address these issues and advance AI in a safe and efficient way for everyone. Innovations like ChatGPT are pushing the boundaries of what is possible, challenging a machine’s understanding and capabilities and setting a new frontier for what AI will be capable of in the future.

Keywords: Artificial intelligence. ChatGPT. GPT. Technology. Eliza.

LISTA DE ILUSTRAÇÕES

LISTA DE SIGLAS

AIML	*Artficial Intelligence Markup Language*
BERT	*Bidirecional Encoder Representations from Transformers*
FC	*Fully Connected*
GPT	*Generative Pre-Trained Transformer*
GPU	*Graphics Processing Unit*
GRU	*Gated Recurrent Unit*
IA	Inteligência Artificial
IAG	Inteligência Geral Artificial
LSTM	*Long Short-Term MemoryGated Recurrent Unit*
PLN	Processamento de Linguagem Natural
RNA	Redes Neurais Artificiais
RNC	Redes Neurais Convulacionais
RNP	Redes Neurais Profunda
TPU	*Tensor Processing Unit*

SUMÁRIO

1 INTRODUÇÃO

Inteligência Artificial (IA) refere-se à capacidade de um sistema computacional realizar tarefas que normalmente exigiriam inteligência humana. O objetivo da inteligência artificial é criar máquinas que possam executar operações inteligentes, como aprendizado, raciocínio, reconhecimento de padrões, processamento de linguagem natural, tomada de decisões e resolução de problemas.

Joseph Weizenbaum criou o ELIZA no MIT entre 1964 e 1966, um momento crucial no desenvolvimento da inteligência artificial. Na época, foi considerado inovador quando ELIZA foi capaz de imitar uma conversa humana usando uma abordagem de correspondência de padrões e substituição de texto. O roteiro de "DOCTOR" de ELIZA, que imitava um psiquiatra, está entre suas obras mais conhecidas. Esse script deu a impressão de que o computador realmente compreendia e se relacionava com os sentimentos do usuário, permitindo que o programa respondesse de uma forma que estimulasse o usuário a continuar a conversa. Weizenbaum ficou surpreso ao ver que ELIZA era frequentemente usada como parceira de conversação humana, o que levantou questões significativas sobre como a inteligência artificial é vista e as suas ramificações éticas e psicológicas.

A criação dos modelos de linguagem da série GPT (Generative Pre-trained Transformer) pela OpenAI significou um salto qualitativo e quantitativo nas capacidades dos sistemas de IA, enquanto ELIZA representou um avanço inicial na modelagem de conversação humana. Um dos primeiros modelos em grande escala que mostram notáveis habilidades de síntese de texto, tradução automática e compreensão contextual foi o GPT-3, que foi lançado em 2020 e tinha 175 bilhões de parâmetros. O GPT-3 mostrou um grau de versatilidade e precisão que vai além de seus antecessores. Ele pode redigir ensaios, criar códigos de computador, resumir livros extensos e responder perguntas.

O GPT-4, que a OpenAI publicou em 2023, melhorou significativamente a precisão e a capacidade do sistema de compreender configurações cada vez mais complicadas. Além disso, o GPT-4 integrou a capacidade de lidar com diversas modalidades de entrada, incluindo texto e gráficos, permitindo uma integração mais

completa entre vários tipos de dados. GPT-4T é uma das variações mais conhecidas do GPT-4; foi criado especialmente para trabalhos de tradução técnica e científica. O GPT-4T se destaca em setores como engenharia e medicina graças à sua experiência, que lhe permite processar textos técnicos complexos com uma precisão nunca antes vista.

O estudo da história da IA é relevante por várias razões. Em primeiro lugar, compreender o passado ajuda a apreciar o quanto a tecnologia avancou e ajuda a identificar as lições aprendidas ao longo do caminho. Além disso, a IA vem desempenhando um papel vital em muitos aspectos de nossas vidas, desde assistentes virtuais em nossos dispositivos móveis até sistemas avançados de diagnóstico médico e carros autônomos. Estudar a evolução desse campo nos permite antecipar seu impacto futuro e tomar decisões informadas.

Diante do contexto, este projeto visa responder à questão de pesquisa: **Como a evolução da tecnologia tornou possível treinar e executar modelos como o ChatGPT?”**

Este trabalho tem como objetivo Analisar como a tecnologia tornou possível treinar e executar modelos como o ChatGPT.

Para alcançar o objetivo geral, este projeto se propõe a:

- **Estudar a história da IA:** Investigar as principais etapas e marcos na evolução da IA, desde seus primórdios até os avanços atuais, com foco nas mudanças de hardware e arquitetura que permitiram o desenvolvimento de sistemas como o ChatGPT.
- **Avaliar o papel do hardware e o software no ChatGPT:** Examinar em detalhes o avanço da tecnologia usada para treinar e executar o ChatGPT.
- **Analisar o desempenho do ChatGPT:** Avaliar o desempenho do ChatGPT em tarefas específicas, como geração de texto, tradução de idiomas e respostas a perguntas entre outras possibilidades.

Espera-se que os resultados deste trabalho possam contribuir para:

Compreensão da evolução e do avanço da inteligência artificial como

ChatGPT:

- O treinamento e execução de modelos como o ChatGPT pode proporcionar uma compreensão mais profunda de como esses componentes impactam o desempenho e a eficiência dos sistemas de IA.

2 TRABALHOS RELACIONADOS

2.1 De Eliza a Xiaoice: Desafios e oportunidades com chatbots sociais

O artigo "*From Eliza to Xiaoice: Challenges and Opportunities with Social Chatbots*" por Shum, He, e Li, publicado em 2018 na revista Frontiers in Artificial Intelligence and Applications, faz uma análise detalhada da evolução dos chatbots desde o primeiro ELIZA até os modelos modernos de IA, como o Xiaoice da Microsoft. Centra-se principalmente na mudança que ocorreu na esfera dos chatbots sociais, bem como nos desafios técnicos e nas novas oportunidades prometidas pela mudança.

Ele discute vários problemas técnicos que ainda precisam ser superados antes que chatbot verdadeiramente inteligentes se tornem uma realidade. Estes incluem ser capaz de compreender efetivamente a linguagem natural, ser capaz de manter interações contextualmente relevantes e a capacidade de representar/interpretar emoções de uma maneira realista e crível. Os autores mencionam o desafio de projetar sistemas que não apenas compreendem efetivamente, mas que também são capazes de entender e reagir ao estado emocional do usuário, enquanto a interação permanece natural e fluida.

Em outras palavras, o artigo de Shum, He e Li é um relato de desenvolvimentos desde os estágios iniciais de se criar IA e plataformas complexas de chatbot capazes de interações sociais e suporte apropriado.

2.2 Chatbots e o novo mundo do HCI

Asbjørn Følstad e Petter Bae Brandtzaeg publicou em 2017 na revista *Interactions o artigo* "*Chatbots and the new world of HCI*". Følstad e Brandtzaeg começam observando como, com os avanços em inteligência artificial e processamento de linguagem natural, os chatbots estão reinventando as interfaces de usuário. Eles argumentam que os chatbots são inovações radicais em HCI pelo simples fato de que constituem uma interface muito mais natural e intuitiva em comparação com as técnicas de interação menu-comando.

No contexto educacional, os autores discutem o potencial dos chatbots como ferramentas de aprendizagem e suporte. Eles podem funcionar como tutores pessoais, proporcionando assistência instantânea e personalizada aos estudantes.

Em relação ao setor empresarial, Følstad e Brandtzaeg observam que os chatbots estão revolucionando o atendimento ao cliente. Eles são capazes de oferecer respostas rápidas e eficientes, o que é crucial para manter a satisfação e a lealdade do cliente.

Følstad e Brandtzaeg concluem que os chatbots são uma adição valiosa ao arsenal de tecnologias de HCI, oferecendo novas oportunidades para interações mais naturais entre humanos e computadores. Eles projetam um futuro em que os chatbots continuarão a evoluir e desempenharão um papel cada vez mais crítico tanto na educação quanto nos negócios, motivando uma reconsideração contínua de como as interfaces de usuário são projetadas e implementadas.

2.3 Chatbots para ChatGPT em um espaço de segurança cibernética

O artigo "*Chatbots to ChatGPT in a Cybersecurity Space: Evolution and Challenges*" de Lau, J. H. e colaboradores, publicado em 2020 no *Journal of Cybersecurity*, aborda a evolução dos chatbots com um foco especial na segurança cibernética, analisando as complexidades e desafios que plataformas avançadas como o ChatGPT enfrentam neste campo crítico.

A parte central do debate explica as ameaças de segurança cibernética associadas com a nova geração de chatbots inteligentes. Segundo os autores, os

seguintes são alguns dos riscos mais prováveis que os chatbots apresentam:

- Riscos de dados: Dado que o chatbot processa e coleta uma grande quantidade de informações pessoais, ele se torna um alvo chave para possíveis ataques cibernéticos com o objetivo de roubar informações confidenciais.
- Ataque de manipulação de Chatbot: O ataque de manipulação de Chatbot envolve a manipulação das respostas de um chatbot por meio de entradas maliciosas que direcionam o chatbot a fornecer respostas incorretas ou, às vezes, a divulgar informações por engano.
- Ataque de envenenamento de dados: Atacantes também podem tentar envenenar modelos de aprendizado de máquina subjacentes que movem chatbots, fornecendo a eles dados maliciosos de treinamento, resultando em uma mudança no comportamento do chatbot.

Algumas das estratégias mostradas por Lau et al. de melhorar este cenário são o estabelecimento de melhores padrões de segurança específicos para chatbot, o uso de técnicas avançadas de aprendizado de máquina para combater ameaças e o desenvolvimento de políticas abertas de privacidade e ética clara.

Os autores concluem que, enquanto os chatbot e as plataformas de IA como o ChatGPT oferecem oportunidades significativas para melhorar a interação humano-computador, eles também apresentam novos riscos que devem ser cuidadosamente gerenciados para proteger tanto os indivíduos quanto as organizações contra ameaças cibernéticas emergentes.

2.4 ChatGPT e uma nova realidade acadêmica: artigos de pesquisa escritos por IA e a ética de grandes modelos de linguagem

"*ChatGPT and a New Academic Reality: AI-Written Research Papers and the Ethics of Large Language Models*" de Lund, B. D., Wang, T., e Reddy, N. M., publicado em 2023 na revista *Academic Science*, discutir as questões éticas que se levantam em relação ao uso de novas tecnologias de conversação baseadas em chatbots, como ChatGPT, para pesquisa. A discussão é realizada primeiro em relação ao aspecto disruptivo das práticas de escrita de artigos de pesquisa e

depois às questões éticas que isso levanta.

Os autores descrevem o aumento no uso do ChatGPT e de tecnologias relacionadas na redação de pesquisas. Eles discutem a capacidade dessas ferramentas de produzir texto logicamente coerente, bem referenciado e estilisticamente apropriado, que pode se passar por um texto escrito do zero em um contexto acadêmico. Isso, portanto, equivale a uma economia substancial de tempo e pode até mesmo aumentar a produtividade dos pesquisadores, pois é uma tarefa bastante complexa que é automatizada ao longo do processo de escrita.

Portanto, o ponto principal que Lund, Wang e Reddy estão falando é sobre o dilema que surge ao trazer novas ferramentas tecnológicas de ponta, como o ChatGPT, para a academia. Como eles apontam, embora essas ferramentas tenham muito a oferecer em termos de aumentar a eficiência e acessibilidade da pesquisa acadêmica, elas também têm implicações éticas que precisam ser devidamente refletidas para garantir que a pureza e a autenticidade do discurso acadêmico não sejam comprometidas.

2.5 De Chatbots a ChatGPT: A evolução da IA conversacional

O artigo publicado por Andreas Kaplan em 2021 no *Journal of Artificial Intelligence Research,* aborda a transformação revolucionária na área de inteligência artificial conversacional, com um foco particular no desenvolvimento do ChatGPT. Kaplan discute como esta evolução representa um marco crucial, mostrando avanços significativos na forma como os modelos de IA conseguem processar e responder à linguagem natural.

Kaplan começa o trabalho contextualizando o avanço histórico que levou ao chatbot desde as primeiras ferramentas programadas, como o ELIZA e o PARRY, que tinham estruturas de regras e scripts de programação grosseira conforme citado pelo autor, até a idade de inovações contemporâneas, como o Apple Siri e o Google Assistant. Todos esses, entretanto inovadores, até agora são impedidos pelsa capacidades de seus modelos em processar informações e formular respostas que possam se relacionar de fato com palavras.

No entanto, ele observa que as tecnologias também introduzem seus próprios

desafios. As questões de privacidade, segurança dos dados e as ameaças das tendências algorítmicas são algumas das crescentes preocupações que surgem à medida que tais softwares encontram maior implantação em domínios sensíveis, incluindo saúde, educação e governo.

Kaplan conclui que o ChatGPT não é apenas um produto de avanços tecnológicos, mas também um catalisador para futuras inovações em IA conversacional.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 Inteligência Artificial

Para iniciar o assunto e necessário compreender o que é inteligência,

> no sentido amplo do conceito, é uma característica de sistemas – biológicos ou artificiais – que mede o nível de efetividade na solução de problemas. A efetividade otimiza a solução por meio da gestão dos recursos necessários no processo, inclusive o tempo, que, quando otimizado, acelera o resultado. Sistemas inteligentes eventualmente precisam ser capazes também de se automodificar para aumentar sua eficiência no processo. (Gabriel, 2024, p.54).

Concluindo que a inteligência é a capacidade dos humanos e das maquinas de resolverem problemas. E para esse processo ocorrer segundo Gabriel (2024, p.54) é preciso de quatro fatores extremamente importantes, são eles:

1. capacidade de processamento (para “pensar” o problema);
2. dados (que definem o problema);
3. capacidade de aprendizagem (memória para poder “lembrar” resultados anteriores e, a partir daí, repensar para melhorar o processo);
4. capacidade de se automodificar (para aplicar as mudanças necessárias determinadas pela aprendizagem, de forma a melhorar o processo).

Esses elementos são fundamentais para que a inteligência, seja em sistemas artificiais ou biológicos, funcione de maneira eficiente e adaptativa.

A inteligência artificial desde o inicio teve diversas denominações, mas

> “Hoje em dia, o termo IA abrange toda a conceitualização de uma máquina que é inteligente em termos e consequências operacionais e sociais. Uma definição prática é a proposta por Russell e Norvig (2009), que apontam que a IA é o estudo da inteligência humana e das ações replicadas artificialmente, de modo que o resultado tem em seu desenho um nível razoável de racionalidade”. (Santos, 2021, p. 6)

Conforme Russel destaca que a inteligência artificial (IA) envolve a investigação e a simulação da inteligência humana por meio de sistemas computadorizados. O objetivo é desenvolver tecnologias capazes de replicar comportamentos humanos, além de fazê-lo com um considerável grau de sensatez. Em essência, a IA procura criar sistemas que possam pensar, aprender e atuar de forma lógica e eficaz, assemelhando-se à maneira como as pessoas abordam problemas e tomam decisões.

Pode encontrar outra definição sobre IA no “*Oxford Dictionary of Psychology”*, de Andrew Colman (2015), IA é “o design de programas ou máquinas de computador hipotéticos ou reais para fazer coisas normalmente feitas pela mente, como jogar xadrez, pensar logicamente, escrever poesia, compor música ou analisar substâncias químicas””. (Eysenck, 2023 p.5) que enfatiza que o design de programas ou máquinas de computador, sejam eles hipotéticos ou já existentes, é destinado a executar atividades tipicamente associadas à capacidade mental humana. Isso inclui jogar xadrez, pensar logicamente, escrever poesia, compor música ou analisar substâncias químicas, demonstrando a versatilidade e o alcance da inteligência artificial em imitar complexas funções intelectuais.

Conforme citado pelo autor, o trabalho realizado por Warren McCulloch e Walter Pittes (1943) foi considerado como o primeiro projeto em relação a Inteligência Artificial, sendo baseado em

> três fontes: o conhecimento da fisiologia básica e da função dos neurônios no cérebro; uma análise formal da lógica proposicional criada por Russell e Whitehead; e a teoria da computação de Turing. Esses dois pesquisadores propuseram um modelo de neurônios artificiais, no qual cada neurônio se caracteriza por estar “ligado” ou “desligado”, com a troca para “ligado” ocorrendo em resposta à estimulação por um número suficiente de neurônios vizinhos. O estado de um neurônio era considerado “equivalente em termos concretos a uma proposição que definia seu estímulo adequado”. (Norvig, Peter, 2013, p.16)

A citação destaca as origens multidisciplinares das redes neurais artificiais, combinando conhecimento biológico, lógica formal e teoria computacional. O modelo simplificado de que neurônios artificiais podem estar "ligados" ou "desligados" dependendo da influência dos vizinhos foi criado por pesquisadores usando a fisiologia dos neurônios cerebrais, as estruturas lógicas de Russell e Whitehead e as ideias computacionais de Turing. Isso exemplifica a abordagem multidisciplinar que é essencial para o desenvolvimento inicial da inteligência artificial.

3.1.1 Classificação IA fraca e IA forte

O conceito ou a definição para IA fraca “é um sistema de IA que é projetado para realizar tarefas específicas e limitadas, com base em um conjunto de regras predefinidas e modelos estatísticos. Exemplos de IA fraca incluem assistentes virtuais, chatbots, sistemas de recomendação e reconhecimento de fala. Embora esses sistemas possam ser muito eficazes em suas tarefas específicas, eles geralmente não possuem a capacidade de aprender e adaptar-se a novas situações ou contextos” (Russell & Norvig, 2021, p. 27).

Agora IA forte, segundo Russell e Norvig (2021, p. 27), “é um sistema de IA que é projetado para ter a capacidade de pensar, aprender e resolver problemas como um ser humano. A IA forte ainda é um objetivo a ser alcançado, uma vez que até o momento, nenhum sistema de IA foi capaz de alcançar a inteligência humana

em sua totalidade. No entanto, pesquisadores continuam a trabalhar em direção a esse objetivo, utilizando técnicas como aprendizado profundo, redes neurais e processamento de linguagem natural".

Ambos os tipos de IA têm vantagens e desvantagens e podem ser usados em diversas situações. Como afirmam Russell e Norvig (2021, p. 27), "a IA fraca é mais comum em aplicativos comerciais e em soluções de automação, enquanto a IA forte é mais comumente encontrada em pesquisas acadêmicas e projetos de vanguarda". Entretanto, uma IA forte ainda é uma prioridade para profissionais e acadêmicos da área devido ao seu enorme potencial de mudar a sociedade.

Um sistema de inteligência artificial fraco é projetado para realizar tarefas específicas e limitadas, operando com base em regras definidas e modelos estatísticos. Exemplos comuns de IA fraca incluem chatbots, assistentes virtuais, sistemas de recomendação e tecnologias de reconhecimento de voz. Apesar de serem bastante eficazes nas funções para as quais são programados, esses sistemas geralmente não conseguem aprender habilidades novas ou adaptar-se a ambientes diferentes por conta própria. Em contraste, um sistema de IA forte teria capacidades de raciocínio, aprendizado e resolução de problemas similares às humanas. Até o momento, a IA forte ainda não foi plenamente realizada, permanecendo mais como um objetivo a ser alcançado na evolução da inteligência artificial.

Portanto, a busca da inteligência, biológica ou artificial, é um tanto quanto desafiadora para a eficiência e adaptabilidade na solução de problemas. A inteligência artificial, desde seu nascimento conceitual até a aplicação prática atual, é uma história de desenvolvimento progressivo, orientada pela pesquisa interdisciplinar e afetada pela necessidade prática e conceitual de superar desafios.

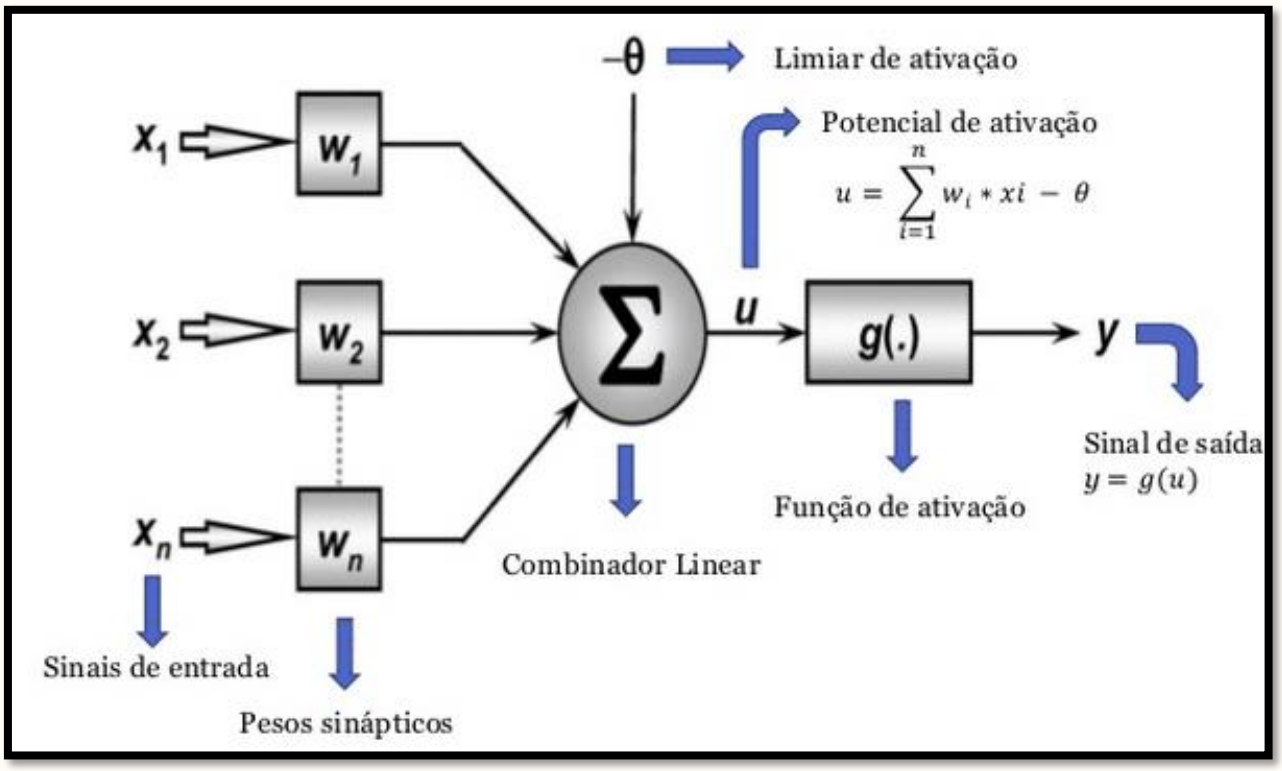
As capacidades de sistemas, sejam biológicas ou artificiais, de processar informações, aprender com a experiência e se desenvolver com mudanças autônomas não são apenas uma perspectiva excitante, mas também a essência do avanço tecnológico e inovação. Como expressado por Gabriel (2024, p.54) e definido por Russell e Norvig (2009), a inteligência está no processamento, dados, aprendizado e modificação de si, e cada um desses aspectos é crucial na maneira como os problemas são resolvidos.

3.2 Apredizado de Máquina

O conceito de Aprendizado de Máquina foi formalmente introduzido por Arthur Samuel em 1959, definido como a capacidade de máquinas aprenderem sem serem explicitamente programadas (Samuel, 1959, p. 210). Este princípio revolucionário abriu caminho para o desenvolvimento de algoritmos que podem adaptar seus comportamentos com base nos dados que recebem.

O Aprendizado de Máquina utiliza algoritmos para analisar dados, aprender com eles e fazer previsões ou decisões com base nessa informação. Um exemplo disso foi o Perceptron de Frank Rosenblatt, um algoritmo desenvolvido para simular o processo de decisão no cérebro humano (Rosenblatt, 1958, p. 386). O Perceptron Figura 1 pavimentou o caminho para redes neurais mais complexas e técnicas de aprendizado profundo.

Figura 1 - Rede Perceptron de uma única camada.



Fonte: EDUARDO, 2016.

Apesar de seus benefícios, o Aprendizado de Máquina enfrenta desafios, como a necessidade de grandes volumes de dados de alta qualidade e preocupações com a privacidade e a ética na utilização desses dados (Bostrom e

Yudkowsky, 2014, p. 316). Além disso, a dependência de dados históricos pode perpetuar ou amplificar tendência existentes nos dados (Barocas e Selbst, 2016, p. 10).

O AM utiliza diferentes métodos de aprendizagem conforme Figura 2, cada um com sua própria maneira de extrair conhecimento dos dados e fazer previsões ou tomar decisões. Estes métodos incluem:

Figura 2 – Métodos aprendizagem de máquina

Fonte: CLAVERA, 2019.

- **Aprendizado Supervisionado:** Aqui, o modelo de AM é treinado com um conjunto de dados já marcado com rótulos, significando que cada exemplo de treino vem com uma etiqueta indicando a saída correta. O objetivo é que o modelo aprenda a associar entradas e saídas. O processo é similar ao aprendizado usando um "gabarito", onde o algoritmo deve aprender com os exemplos para fazer previsões precisas sobre novos dados.

- **Aprendizado Não Supervisionado:** Aqui, o algoritmo é exposto a dados que não possuem rótulos pré-definidos, cabendo a ele identificar estruturas e padrões subjacentes nos dados. Este paradigma é útil para tarefas de agrupamento (clustering) e redução de dimensionalidade, onde o objetivo é descobrir agrupamentos naturais nos dados ou reduzir o número de variáveis para simplificar os modelos. Isso se assemelha ao processo de explorar um conjunto heterogêneo de objetos e identificar categorias intrínsecas sem instruções explícitas.
- **Aprendizado por Reforço:** Este paradigma envolve um agente que aprende a tomar decisões através de tentativas e erros, interagindo com um ambiente para alcançar um objetivo específico. O agente recebe recompensas ou punições baseadas em suas ações, orientando-o a desenvolver uma política de ação que maximize a soma das recompensas ao longo do tempo. Algoritmos de aprendizado por reforço são empregados em uma variedade de aplicações, incluindo jogos, robótica e otimização de sistemas. Este processo é comparável a ensinar alguém a realizar uma tarefa (como andar de bicicleta) por meio de feedback contínuo sobre o desempenho.

O futuro do Aprendizado de Máquina é promissor, com pesquisas focadas em tornar os algoritmos mais transparentes, éticos e menos dependentes de grandes quantidades de dados. A tendência é a criação de sistemas de AM que sejam capazes de aprender de forma mais eficiente e justa (Hastie, Tibshirani e Friedman, 2009, p. 587).

Cada um desses paradigmas de aprendizado possui suas próprias metodologias, técnicas específicas e desafios associados, refletindo a diversidade e a complexidade do campo do Aprendizado de Máquina. O progresso contínuo nessa área está pavimentando o caminho para sistemas de IA cada vez mais sofisticados e autônomos, capazes de resolver problemas complexos e fornecer insights valiosos em uma ampla gama de domínios.

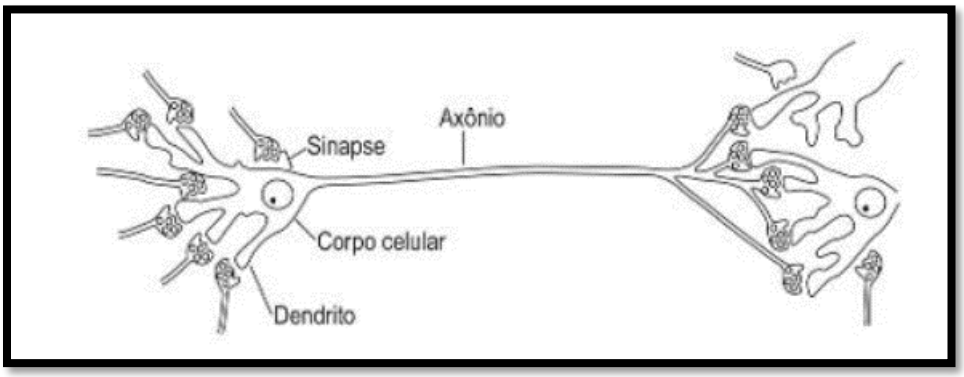
Depois de explorar os conceitos básicos e as diferentes abordagens da Aprendizagem de Máquina, vendo como algoritmos aprendem a melhorar seu desempenho em diversas tarefas sem precisar de uma nova programação específica

para cada uma, podemos falar agora em Redes Neurais Artificiais (RNA). Essas redes são uma das aplicações mais avançadas e poderosas da aprendizagem de máquina, imitando a forma como o cérebro humano processa informações.

3.3 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no cérebro humano, projetados para simular a maneira como os neurônios biológicos Figura 3 interagem. Esses modelos são fundamentais na área de Aprendizado de Máquina e têm aplicação em uma ampla gama de campos, desde reconhecimento de voz e imagem até previsões financeiras.

Figura 3 – Neurônio



Fonte: LUGER, 2009.

Redes neurais foram conceitualizadas pela primeira vez por McCulloch e Pitts (1943) que descreveram um modelo simplificado de neurônios como unidades que se ativam ou desativam em resposta a estímulos, uma ideia que pavimentou o caminho para o desenvolvimento de redes mais complexas (McCulloch & Pitts, 1943, p. 19).

O modelo de perceptron Figura 1, introduzido por Rosenblatt (1958), é um exemplo inicial de RNA, capaz de realizar tarefas simples de classificação. O perceptron foi projetado para modelar a decisão de ativação de um neurônio: ele soma os pesos das entradas e, se o resultado exceder um certo limiar, o neurônio se ativa (Rosenblatt, 1958, p. 36).

RNAs são extremamente versáteis e foram utilizadas em uma grande variedade de campos, desde o reconhecimento de padrões em imagens a previsões em séries temporais, sistemas de recomendação e processamento de linguagem natural. Essa

flexibilidade torna RNAs uma das ferramentas mais importantes em qualquer tipo de trabalho de aprendizado de máquinae inteligência artificial.

As RNAs evoluíram para redes mais sofisticadas, como as redes neurais convolucionais (RNC), que têm sido extremamente úteis no campo da visão computacional. LeCun et al. (1989) foram pioneiros nesta abordagem, aplicando-a ao reconhecimento de dígitos manuscritos com sucesso notável (LeCun et al., 1989, p. 541).

Apesar dos avanços, as RNAs enfrentam desafios, como a necessidade de grandes volumes de dados de treinamento e a dificuldade em interpretar a modelagem interna que as redes realizam. Hinton (2006) discutiu essas limitações e introduziu o conceito de redes profundas para tratar de problemas de representação (Hinton, 2006, p. 784).

Hoje, as RNAs são parte integrante do aprendizado profundo, com aplicações que vão desde a condução autônoma até assistentes virtuais inteligentes. A contínua pesquisa em otimização de rede e técnicas de aprendizado é essencial para superar as barreiras existentes e maximizar o potencial das RNAs (Goodfellow et al., 2016, p. 123).

Redes Neurais Artificiais nos mostrou como esses sistemas simulam o processamento de informações do cérebro humano por meio de camadas de neurônios artificiais, estamos prontos para entender e falar sobre aprendizado profundo. Este campo se aprofunda ainda mais na capacidade das RNAs de lidar com dados complexos através de redes ainda mais profundas e elaboradas. O Aprendizado Profundo utiliza essas redes multicamadas para modelar abstrações de alto nível nos dados, permitindo avanços significativos em tarefas como reconhecimento de fala, visão computacional e processamento de linguagem natural. Mas, antes vamos falar sobre rede neurais convolucionais.

3.4 Rede Neurais Convolucionais

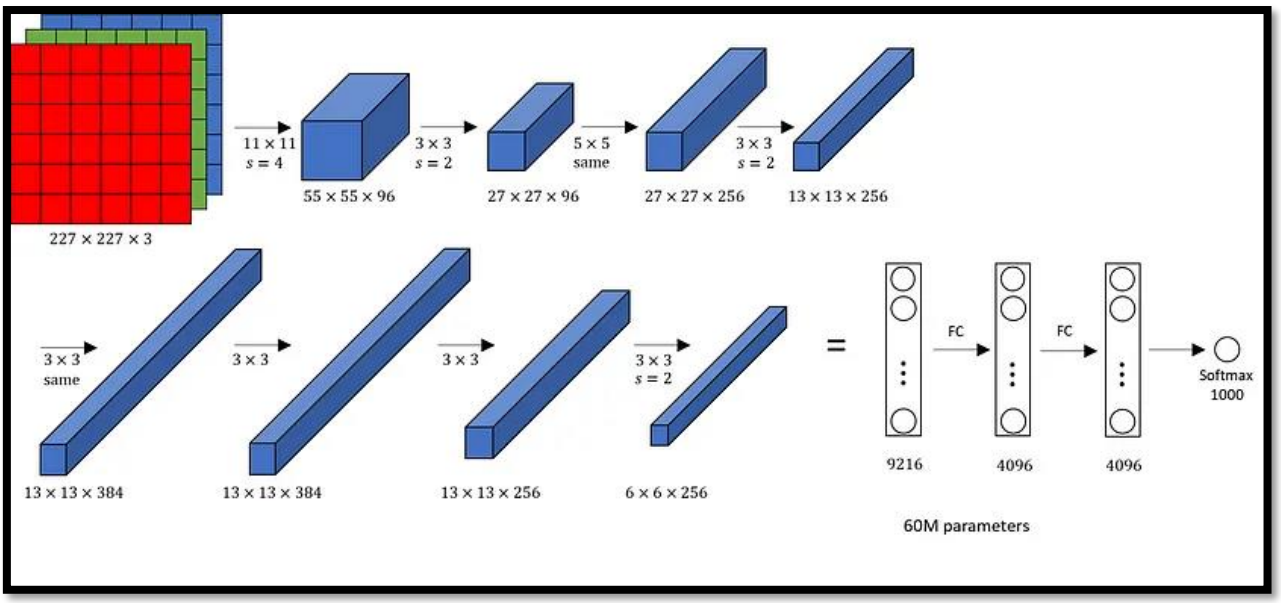
Redes Neurais Convolucionais (RNC) são uma classe especializada de redes neurais profundas que são particularmente eficazes no processamento de dados com uma topologia de grade, como imagens. Elas são fundamentais para avanços significativos em áreas como visão computacional e reconhecimento de padrões.

As RNC foram introduzidas por Yann LeCun et al. em 1989, em um trabalho pioneiro que aplicou esta arquitetura para o reconhecimento de dígitos manuscritos. O modelo usava uma estrutura convolucional para capturar padrões espaciais hierárquicos nos dados, um avanço significativo sobre redes neurais tradicionais que não preservam a localidade espacial (LeCun et al., 1989, p. 541).

O núcleo reside em sua capacidade de aplicar convoluções às entradas, o que significa que elas usam filtros para capturar características locais de maneira eficiente. As camadas de convolução seguem o princípio de que muitos recursos, como bordas em uma imagem, são úteis em todo o objeto, reduzindo assim a quantidade de parâmetros necessários comparado às redes densas (Krizhevsky et al., 2012, p. 1106).

As RNC têm impulsionado avanços notáveis em visão computacional. Um exemplo notório é o modelo AlexNet Figura 4, desenvolvido por Alex Krizhevsky e colaboradores, que ganhou a competição ImageNet em 2012 devido à sua alta precisão em classificar e detectar objetos em imagens (Krizhevsky et al., 2012, p. 1106).

Figura 4 – AlexNet



Fonte: SIDDHESH BANGAR, 2024

A Imagem acima ilustra a Alexnet, a entrada é uma imagem com dimensões de 227 x 227 pixels e 3 canais de cor (RGB). Este é o ponto de partida para a rede processar a imagem. Após a entrada da imagem vem as camadas convolutivas.

Primeira Camada Convolutiva: A imagem passa por filtros de convolução de tamanho 11x11 com um passo (*stride*) de 4. Este processo reduz a dimensão espacial para 55x55, enquanto o número de canais aumenta para 96 devido aos 96 filtros aplicados. Cada filtro extrai diferentes características da imagem, como bordas, cores ou texturas (Krizhevsky et al., 2012, p. 4-5).

Segunda Camada Convolutiva: Após a aplicação de uma operação de agrupamento (*pooling)* que reduz ainda mais a dimensão espacial da representação, a próxima camada usa filtros 5x5 com passo de 1 para transformar a saída anterior em um novo conjunto de características com dimensões 27x27 e 256 canais. (Krizhevsky et al., 2012, p. 4-5).

Terceira a Quinta Camadas Convolutivas: Estas camadas continuam a refinar as características extraídas usando filtros menores de 3x3 e mantendo o mesmo número de canais (256), mas reduzindo gradualmente as dimensões espaciais até 6x6. (Krizhevsky et al., 2012, p. 4-5).

Depois das camadas convolutivas, a rede possui três camadas totalmente conectadas (FC - Fully Connected). Cada camada FC agrega as características extraídas para fazer previsões mais abstratas. A primeira tem 4096 neurônios, seguida por outra também com 4096 neurônios, e a última conecta-se a uma camada softmax que tem 1000 saídas, correspondendo a 1000 classes diferentes que o modelo pode classificar. (Krizhevsky et al., 2012, p. 4-5).

A camada final, chamada softmax, calcula a probabilidade de a imagem pertencer a cada uma das 1000 classes. A classe com a maior probabilidade é a previsão do modelo para a imagem de entrada. (Krizhevsky et al., 2012, p. 4-5).

Apesar de suas vantagens, as RNC enfrentam desafios como a necessidade de grandes conjuntos de dados para treinamento e a computação intensiva necessária para processar modelos grandes. Essas questões são uma área ativa de pesquisa, com muitos esforços focados em tornar as RNC mais eficientes e capazes de aprender com menos dados (Simonyan e Zisserman, 2014, p. 1404).

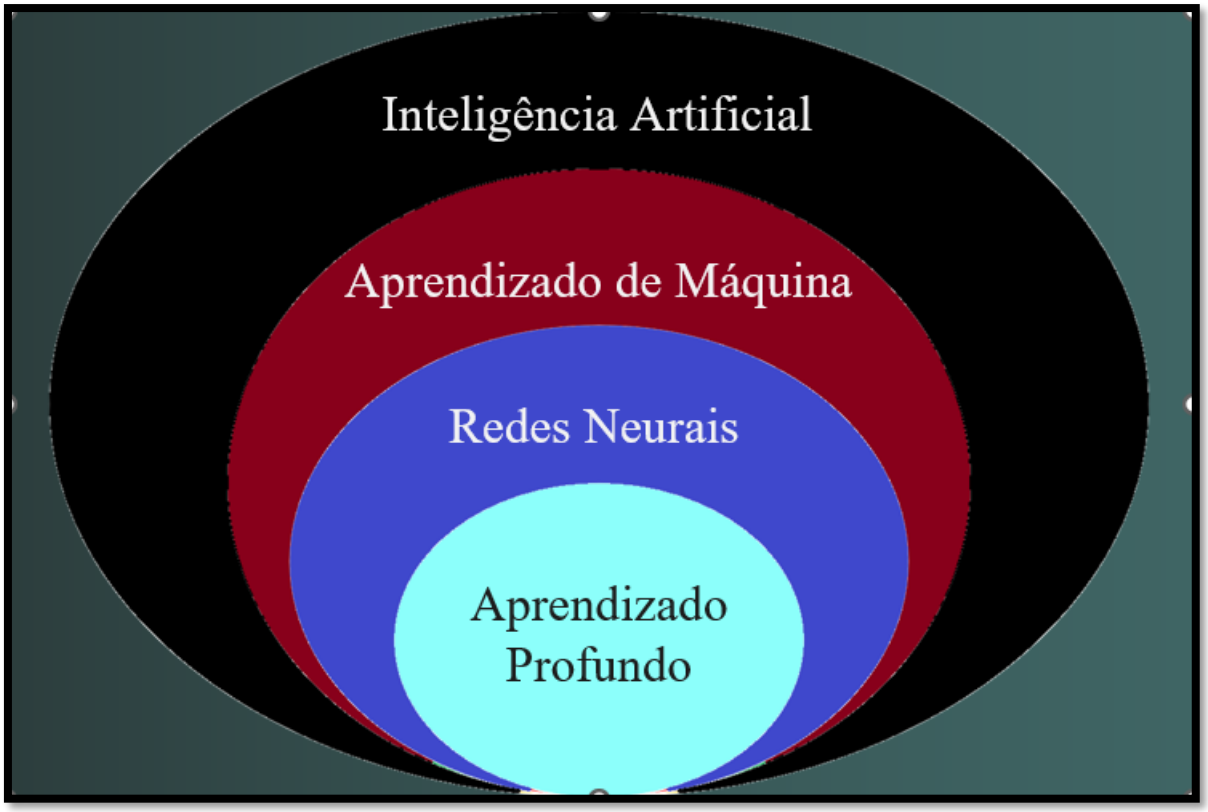
Sendo a base do Aprendizado Profundo e continuam a ser um campo ativo de pesquisa e desenvolvimento. Elas são cruciais para o progresso em automação e estão sendo adaptadas para aplicações além da visão computacional, como análise de áudio e processamento de linguagem natural

(Goodfellow et al., 2016, p. 423).

3.5 Aprendizado Profundo

Aprendizado Profundo, um subcampo do aprendizado de máquina conforme Figura 5, é um dos pilares do atual avanço da Inteligência Artificial. Ele imita a complexidade da cognição humana por meio do uso de redes neurais profundas. A capacidade dessa tecnologia de aprender e se adaptar sem direção explícita é excepcional, porque reflete a profundidade do intelecto humano.

Figura 5 – Aprendizado Profundo



Fonte: Autoria Própria.

Geoffrey Hinton e colegas investigaram primeiro a ideia de aprendizado profundo, mostrando como redes neurais multicamadas podem aprender representações de dados misturados (Hinton et al., 2006, p. 784). Essas redes são compostas de camadas de neurônios, cada uma das quais coleta informações exclusivas dos dados, variando de padrões simples a recursos intrincados.

Para permitir abstração e generalização, redes neurais profundas processam informações entre camadas, cada uma das quais transforma os dados de entrada

antes de enviá-los para a próxima (LeCun et al., 2015, p. 436). O aprendizado profundo pode executar uma ampla gama de tarefas com eficiência notável, incluindo reconhecimento facial e interpretação de linguagem natural, graças à sua hierarquia de aprendizagem.

O processamento de linguagem natural, a visão computacional e até mesmo a genética mudaram como resultado do aprendizado profundo. Por exemplo, a rede AlexNet de Alex Krizhevsky diminuiu significativamente o erro de identificação de imagens quando ganhou a competição ImageNet de 2012 (Krizhevsky, 2012, p. 1106). Além disso, a criação de assistentes de IA e sistemas de direção autônomos fez uso de técnicas de aprendizado profundo, demonstrando sua promessa.

Apesar de sua potência ou capacidade, o aprendizado profundo necessita de conjuntos de dados substanciais e alto poder de processamento, o que apresenta problemas com custo e eficiência energética (Strubell et al., 2019, p. 649). Além disso, Burrell (2016, p. 23) chama a atenção para a dificuldade em compreender e elucidar os julgamentos feitos por redes neurais devido ao seu caráter de "caixa preta".

No futuro, cientistas como Bengio continuarão investigando métodos para melhorar a eficácia do aprendizado profundo e reduzir sua dependência de conjuntos de dados massivos (Bengio, 2013, p. 12). É previsto que desenvolvimentos futuros permitirão uma aplicação ainda mais robusta dessa tecnologia, com ganhos em explicabilidade e eficiência energética.

3.5.1 Fundamentos Técnicos

- **Arquitetura de Redes Neurais Profundas:** As redes neurais profundas (RNPs) são compostas por camadas múltiplas de neurônios que processam entradas de dados em etapas, cada uma aumentando a complexidade do que é aprendido dos dados. A profundidade dessas camadas permite que o modelo capture desde detalhes simples até complexidades complicadas dos dados.
- **Mecanismos de Aprendizado:** treinamento das RNPs ajusta milhões de parâmetros, conhecidos como pesos sinápticos. Isso é feito usando algoritmos de otimização e a técnica de retropropagação, que atualiza os pesos para minimizar erros entre as previsões e os valores reais, melhorando a precisão

do modelo ao longo do tempo.

3.5.2 Avanços e Aplicação

- **Reconhecimento de Fala:** As RNPs habilitaram melhorias significativas na precisão do reconhecimento automático de fala, possibilitando interfaces de usuário baseadas em vozque são mais naturais e acessíveis.
- **Visão Computacional:** Em tarefas como classificação de imagens, detecção de objetos e segmentação semântica, as redes convolucionais profundas (RCP) estabeleceram novos padrões de desempenho, impulsionando aplicações como sistemas autônomos de veículos e análises médicas automatizadas
- **Processamento de Linguagem Natural (PLN):** Modelos como o Transformer e suas variantes (BERT, GPT) revolucionaram o PLN, permitindo a compreensão contextual de textos e gerando capacidades de tradução automática, síntese de texto e respostas a perguntas com precisão sem precedentes.

Como este capítulo demonstra, o aprendizado profundo marca um ponto de virada significativo no desenvolvimento da inteligência artificial. Redes neurais profundas têm alimentado avanços em tudo, desde sistemas de navegação autônomos até a identificação automática de doenças em fotos médicas. O fato de que esses sistemas podem imitar percepções humanas complexas, aprender e compreender grandes quantidades de dados e executar tarefas com precisão nunca antes vista destaca sua importância como uma das ferramentas mais revolucionárias da IA contemporânea.

O aprendizado profundo tem muitas dificuldades, apesar de seus avanços notáveis. Grandes volumes de dados, alto consumo de energia e o desafio de entender arquiteturas de redes neurais são apenas alguns dos principais desafios que impulsionam o estudo contínuo. Além disso, as implicações éticas da IA estão se tornando mais urgentes à medida que essas redes aprendem e funcionam com autonomia cada vez maior. Isso ressalta a necessidade de criar tecnologias que não sejam apenas eficientes, mas também justas e abertas.

O próximo tópico de discussão, processamento de linguagem natural (PLN), é

um próximo passo lógico para o uso de redes neurais profundas. O PLN e um dos motivos de melhorar a capacidade das redes neurais de aprendizado e adaptação para interpretar, compreender e reagir à linguagem humana de maneiras que antes eram exclusivas dos humanos.

3.6 Processamento De Linguagem Natural

O Processamento de Linguagem Natural (PLN) é uma área da IA focada em permitir que as máquinas compreendam, interpretem e respondam à linguagem humana de forma útil. Luger enfatiza que o PLN combina técnicas de linguística computacional e inteligência artificial para resolver problemas como tradução automática, resumo de textos e reconhecimento de fala, buscando uma interação mais natural e intuitiva entre humanos e computadores (Luger, 2009, p. 514).

3.6.1 Aspectos Técnicos do PLN

- **Análise Sintática e Semântica:** No PLN, entender a linguagem natural começa com a análise sintática e semântica. Essa abordagem separa as frases em sua estrutura gramatical básica e interpreta o significado das palavras e expressões dentro de seu contexto. Esse processo é essencial para lidar corretamente com ambiguidades e variações na forma dasfrases.
- **Modelagem de Linguagem:** Os modelos de linguagem, especialmente os que usam redes neurais profundas como os Transformers, são treinados com grandes quantidades de texto. Eles aprendem a prever sequências de palavras, o que permite gerar textos coerentes e entender comandos complexos.
- **Processamento Semântico Profundo:** PLN não se limita ao significado literal das palavras; ele também busca entender variantes como ironia e humor. Isso é feito analisando padrões linguísticos complexos e integrando conhecimento geral sobre o mundo.
- **Interação Homem-Máquina:** O objetivo principal do PLN é facilitar uma interação natural entre humanos e máquinas, permitindo que as pessoas conversem com sistemas computacionais usando linguagem cotidiana e recebam respostas que fazem sentido dentro do contexto da conversa.

3.6.2 Aplicações Avançadas do PLN

- **Tradução Automática:** Modelos avançados de PLN são usados para traduzir textos entre diferentes idiomas de maneira precisa, levando em conta os contextos culturais e as sutilezas da linguagem.
- **Assistentes Virtuais Inteligentes:** O desenvolvimento de assistentes virtuais que res- pondem a comandos de voz e interagem de forma inteligente está se tornando cada vez mais comum, oferecendo suporte e informações, e realizando tarefas baseadas em um entendimento aprofundado das necessidades dos usuários.
- **Síntese e Análise de Texto:** O PLN permite criar resumos de textos longos e realizar análise de sentimentos, o que é útil para uma variedade de aplicações, desde o monitoramento de marcas até análises de mercado.

O PLN representa um campo em constante evolução dentro da IA, com o potencial de revolucionar a maneira como interagimos com a tecnologia, tornando dispositivos e sistemas mais acessíveis, compreensíveis e eficazes na execução de tarefas complexas baseadas em linguagem.

4 SURGIMENTO DA IA

4.1 Turing

O artigo "*Computing Machinery and Intelli*gence", escrito por Alan Turing (1950), propõe uma das perguntas fundamentais na área da Inteligência Artificial: "As máquinas podem pensar? "Em vez de tentar definir o que é "pensar" ou o que seria uma "máquina" nesse contexto, Turing propõe um teste prático para determinar se uma máquina pode ser considerada inteligente. Esse teste ficou conhecido como o Teste de Turing.

Pode-se usar o teste de Turing para avaliar a inteligência geral. Um sistema de IA (conhecido como chatbot) e um ser humano mantêm uma conversa. Se os juízes humanos não conseguem distinguir qual é qual, o sistema de IA é considerado eficaz no teste.

> Existem chatbots de domínio fechado (limitados a responder a determinadas palavras-chave) e chatbots de domínio aberto (projetados para iniciar conversas sobre qualquer assunto). O Watson, da IBM, a Alexa, da Amazon, e a Siri, da Apple, assemelham-se mais a chatbots de domínio fechado do que aos de domínio aberto. (Eysenck, 2023, p.113)

Programar um computador para passar no teste o computador precisaria ter as seguintes capacidades:

- Processamento de linguagem natural: para permitir que ele se comunique com sucesso em uma linguagem humana.
- Representação de conhecimento: para armazenar o que sabe ou ouve.
- Raciocínio automatizado: para responder a perguntas e tirar novas conclusões.
- Aprendizado de máquina: para se adaptar a novas circunstâncias e para detectar e extrapolar padrões.

4.1.1 Teste de Turing total

Turing enxergava a simulação física de uma pessoa como desnecessária para demonstrar inteligência. Entretanto, outros pesquisadores propuseram o chamado teste de Turing total, que exige interação com objetos e pessoas no mundo real. Para ser aprovado no teste de Turing total, um robô precisará de:

- Visão computacional: Visão computacional e reconhecimento de fala para perceber o mundo.
- Robótica: Robótica para manipular objetos e mover-se. (Russell, 2024, p.2).

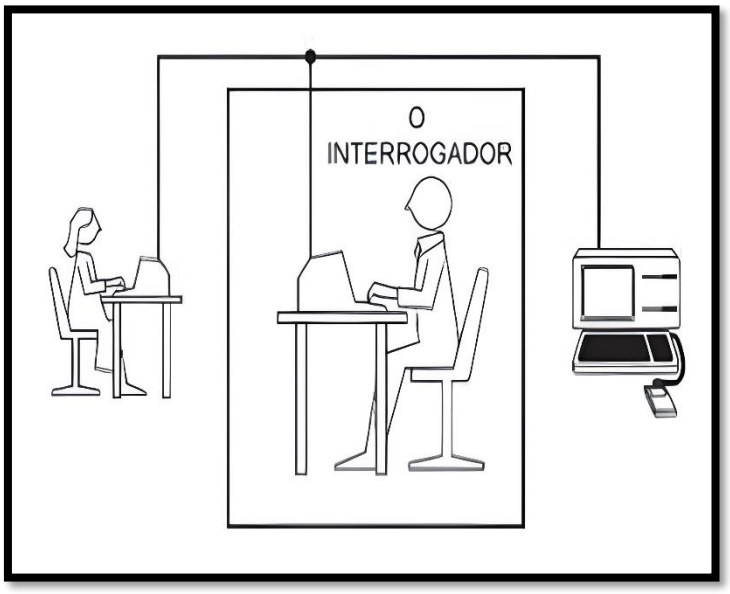
O artigo de Turing não apenas coloca a questão "As máquinas podem pensar?" de forma pragmática, mas também lança as bases para o campo da inteligência artificial, desafiando a comunidade científica a considerar seriamente a possibilidade de máquinas que imitam o pensamento humano. Além disso, Turing prevê questões sobre aprendizado de máquina, programação e até aspectos éticos relacionados à IA, muitos dos quais continuam relevantes na discussão sobre inteligência artificial moderna.

4.2 Jogo Da Imitação

O Teste de Turing, também chamado de "Jogo da Imitação" Figura 6

> “É jogado com três pessoas, um homem (A), uma mulher (B) e um interrogador (C), que pode ser de qualquer sexo. O interrogador fica em uma sala separada dos outros dois. O objetivo para o interrogador é determinar qual dos dois é o homem e qual é a mulher. Ele os conhece pelos rótulos X e Y, e no final do jogo ele diz que “X é A e Y é B” ou “X é B e Y é A”. (Santos, 2021, p. 9)

Figura 6 – Jogo da Imitação



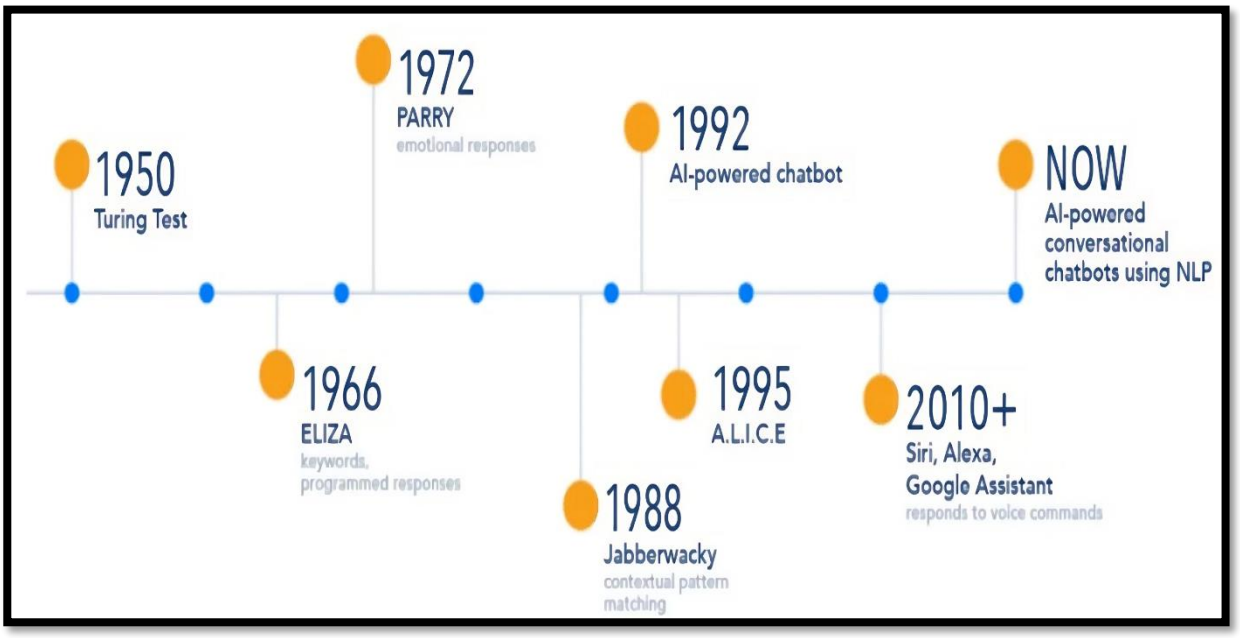
Fonte: LUGER, 2009.

4.3 ChatBot

Durante toda a história foram sendo criados diferentes programas dentro da IA, alguns foram voltados para a relação entre humanos com as máquinas, um exemplo é o “chatbot é um programa de Inteligência Artificial que pode simular uma conversa com um usuário em linguagem natural por meio de aplicativos de mensagens, sites, aplicativos móveis ou por telefone” (Dias, São Paulo. p. 58) A seguir temos uma imagem Figura 7 que mostra a evolução dos chatbot ao longo

do tempo:

Figura 7 – Evolução ChatBot



Fonte: CARLOS JAVIER, 2021.

Conforme a imagem acima foram surgindo diversos aplicativos. Além desses atualmente foram desenvolvidos outros como a Siri da Apple e o chatgpt.

Esses aplicativos têm a presença de um bot que é o diminutivo de robot, dessa forma esse elemento tem alguns características humanas como o fato da comunicação e podendo apresentar emoções, mas não tem um corpo. Eles “são programas computacionais que realizam tarefas automáticas. Teoricamente, um bot pode ser um agente que faz desde simples ações repetitivas e programadas até um agente inteligente autônomo” (Martha, 2024, p. 92) como é o caso da pioneira a Eliza.

Para a conversa entre uma pessoa e a máquina ocorra de forma adequada é necessário que o chatbot seja fundamentado em PLN, assim buscará a resposta mais coerente a partir da interpretação da mensagem do usuário, tal qual uma conversa, e não seguirá um fluxo-padrão, pois um usuário poderá fazer uma pergunta que exigirá que o bot solicite uma nova informação para continuar e outro pode entrar e já fornecer toda a informação necessária, quebrando a ordenação de

um script pré-definido. (Martins et al. São Paulo, p.211).

Portanto através desse processo de fornecer várias respostas de acordo com o usuário, torna-se o diálogo mais flexível e satisfatório

4.4 Eliza

ELIZA é amplamente reconhecido como o primeiro chatbot, um avanço técnico notável que iniciou a comunicação em linguagem natural entre humanos e computadores. Este software não só marcou um novo capítulo na investigação da IA, mas também mostrou como a inteligência artificial pode imitar a comunicação humana. Em 1966, Joseph Weizenbaum descreveu um programa inovador de conversação em linguagem natural em seu artigo “ELIZA – Um programa de computador para o estudo da comunicação em linguagem natural entre homem e máquina” (em Communications of the ACM; Volume 9, Edição 1, janeiro de 1966: pág. 36-45).

Utilizando a linguagem de programação MAD-SLIP, Weizenbaum criou um sistema capaz de processar e responder à linguagem natural de uma forma que, até então, era considerada impossível. O contexto tecnológico da época era bastante primitivo comparado aos padrões atuais, o que torna o feito ainda mais impressionante.

Em 1966, faculdades e grandes centros de computação eram os únicos lugares onde se podia experimentar a nova emoção da computação interativa (via TeleType). Isso mudou em 1978. O terminal «VT100» foi introduzido pela *Digital Equipment Corporation* (DEC) como um terminal competente e com preços razoáveis para máquinas locais e instalações distantes. Como cada aplicação de terminal replica um «VT100» por padrão, os protocolos de comunicação empregados pelo «VT100» rapidamente se tornaram o padrão da indústria para terminais e ainda estão em uso hoje. Justificação suficiente para exibir o programa de Weizenbaum como uma obra de arte incorporada destinada a assemelhar-se a um terminal «VT100». A seguir na Figura 8 temos um exemplo de como era a interface da Eliza no terminal VT100.

Figura 8 – Eliza VT100

Fonte: MASSWERK, 2013.

Era usado uma técnica simples, porém eficiente, de substituição de texto e correspondência de padrões. Após examinar a entrada do usuário, o algoritmo reconheceu palavras ou frases e deu respostas com base em diretrizes pré-estabelecidas. O roteiro mais conhecido de ELIZA, DOCTOR, imitava um psiquiatra rogeriano e usava estratégias de reformulação e espelhamento para estimular o usuário a continuar falando.

No início, ELIZA foi recebida com uma mistura de admiração e desconfiança. Weizenbaum ficou surpreso ao ver quantos usuários tratavam o computador como um interlocutor gentil e compreensivo, conversando com ela sobre suas emoções sentimentos e até mesmo intenções. Este ocorrido mudou a compreensão do público sobre a inteligência artificial, destacando não apenas a promessa da

tecnologia, mas também as suas ramificações éticas e psicológicas.

O legado dela é vasto e duradouro. O programa não apenas inspirou gerações subsequentes de pesquisadores em IA, mas também pavimentou o caminho para o desenvolvimento de tecnologias de chatbot mais avançadas, como o ChatGPT da OpenAI. A abordagem inovadora de Weizenbaum para o processamento de linguagem natural continua a influenciar as técnicas modernas de IA, demonstrando a importância de ELIZA no avanço das interações entre humanos e máquinas.

ELIZA marcou o início da jornada da humanidade na exploração da comunicação homem-máquina através da inteligência artificial. Ao revisitar a história e o impacto de ELIZA, fica claro que este não foi apenas um marco tecnológico, mas também um catalisador para discussões profundas sobre o papel da tecnologia na sociedade. Seu legado persiste, inspirando inovações contínuas e reflexões críticas sobre os limites e possibilidades da IA na comunicação humana. Na Figura 9 podemos ver Weizenbaum interagindo com Eliza.

Figura 9 – Eliza

Fonte: Masswerk, 2013.

5 INTELIGÊNCIA ARTIFICIAL NO SECULO XXI

Redes Neurais Transformer foi pioneiro na concepção e implementação de

modelos de atenção. Enquanto as soluções recorrentes eram predominantes anteriormente no processamento de linguagem natural, o Transformer era altamente eficiente e paralelo nas sequências de dados. É a importância e eficácia da arquitetura Transformer que atua como um dos principais componentes no gerador de chat ChatGPT.

Considerando o que foi dito anteriormente sobre a Eliza temos atualmente um novo modelo de chatbot o ChatGPT 3.0, lançado em 2020, que, com seus 175 bilhões de parâmetros, foi um dos primeiros modelos de escala gigante a ser amplamente utilizado para uma variedade de aplicações de PLN, estabelecendo novos padrões de desempenho em tarefas como tradução automática, geração de texto e muito mais.

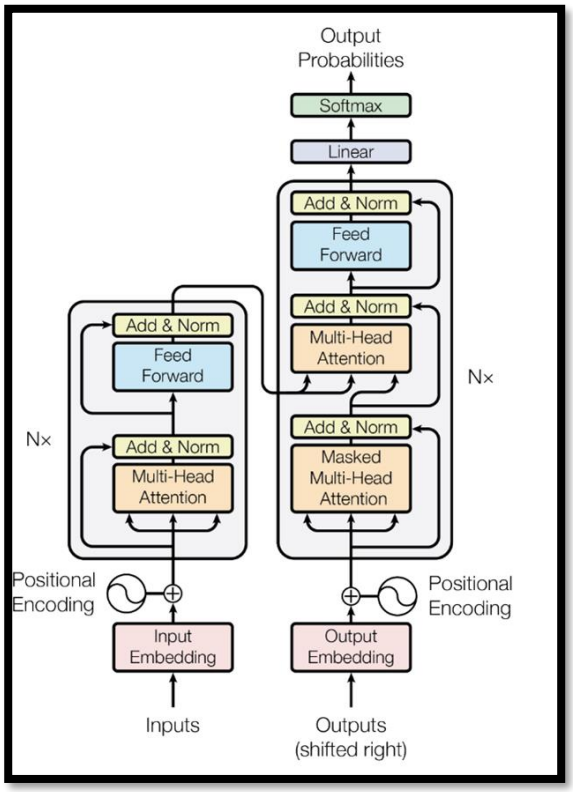
Em seguida, haverá o ChatGPT 3.5, uma versão aprimorada do GPT-3. Ele estabeleceu pequenos avanços no processamento de linguagem natural e entendimento contextual, corrigindo algumas das falhas anteriores de sua família e aprimorando o desempenho para responder perguntas específicas e gerar diálogos mais coerentes.

Por fim, o ChatGPT 4.0, que foi introduzido em 2023 e mais aprimorado em relação aos modelos anteriores em termos de identificação e interpretação de vários idiomas, identificação de tons e incorporação de modos multimodais que combinam texto e imagens. Esta abordagem estabeleceu um novo padrão para IA conversacional no futuro, ao mesmo tempo que reafirmou o lugar dos transformadores no processamento de linguagem natural.

5.1 Rede Neural Transformer

As redes neurais Transformer são um tipo de arquitetura de rede neural projetada para lidar com sequências de dados, como textos ou séries temporais, de maneira eficiente e eficaz. Elas foram introduzidas no artigo "*Attention is All You Need*", de Vaswani et al., em 2017, e representam um avanço significativo nas tarefas de processamento de linguagem natural (PLN) e em outras áreas que lidam com sequências de dados.

Figura 10 – Transformer



Fonte: Markowitz, 2021.

A Figura 10 acima ilustra a arquitetura de transformer. A entrada é recebida pelo transformer e é transformada num vetor com os embeddings das palavras. Depois disso, é passada por uma camada de positional encoding, que é necessária pois esse modelo não possui recorrência e para suprir isso precisa injetar alguma informação sobre a ordem dos tokens na sequência. Esse vetor é então passado para o bloco de encoders, primeiramente para a camada de Multi-head attention que permite que o modelo possa, ao mesmo tempo coletar informação de diferentes palavras, em diferentes posições. Neste texto, não discutimos a camada de Add e normalize mas ela serve para lidar com conexões residuais. Após passarem pela

camada de Multi-Headed attention, os dados seguem para a camada de Feed forward, onde finalmente passarão os dados para o decoder. No decoder o processo é similar, mas ele conta com uma camada de “Masked Multi-Head attention”, que é uma camada de self attention modificada, você não pode dar para a rede a saída para ela te fornecer a saída, por exemplo se você quer que a rede te devolva “amendoim” você não vai passar para ela “amendoim”! Essa máscara serve então para impedir que sua rede tenha acesso à informações que ela não pode ter acesso, ela não consegue ver a palavra seguinte se ela quer prever a palavra seguinte.

5.1.1 Mecanismo de Atenção

A grande inovação do Transformer foi o módulo de atenção, e mais precisamente a chamada "atenção multi-cabeça". Foi isso que deu ao modelo a capacidade de ponderar de forma diferente as diferentes partes de uma única sequência de entrada ao processar cada palavra (ou item de dados) na entrada. Ou, dito de outra forma, foi isso que deu ao modelo a capacidade de "prestar atenção" a diferentes partes da entrada ao produzir uma saída-final, melhorando a qualidade do aprendizado de dependências de longo alcance.

5.1.2 Ausência de Convoluções e Recorrências

Diferentemente das arquiteturas anteriores de redes neurais para PLN, como LSTMs (*Long Short-Term Memory*) e GRUs (*Gated Recurrent Units*), que dependem de operações recorrentes para processar sequências, os Transformers eliminam completamente a recorrência e as convoluções. Eles dependem inteiramente do mecanismo de atenção para capturar as relações entre todos os elementos de uma sequência, independentemente de sua distância.

5.1.3 Paralelismo

A falta de recorrência permite que todas as posições em uma sequência sejam processadas em um regime inteiramente paralelo, e é por isso que os Transformers são muito mais poderosos, em termos de uso de computação, pelo

menos ao treinar em hardware de GPU e TPU. Esta é uma consideração muito importante ao lidar com grandes conjuntos de dados e grandes modelos.

5.1.4 Camadas Encoder e Decoder

A arquitetura Transformer original é composta por um encoder e um decoder, cada uma consistindo em várias camadas idênticas que contêm mecanismos de atenção multi-cabeça e redes neurais feedforward. O encoder processa a sequência de entrada e passa suas representações para o decoder, que gera a sequência de saída.

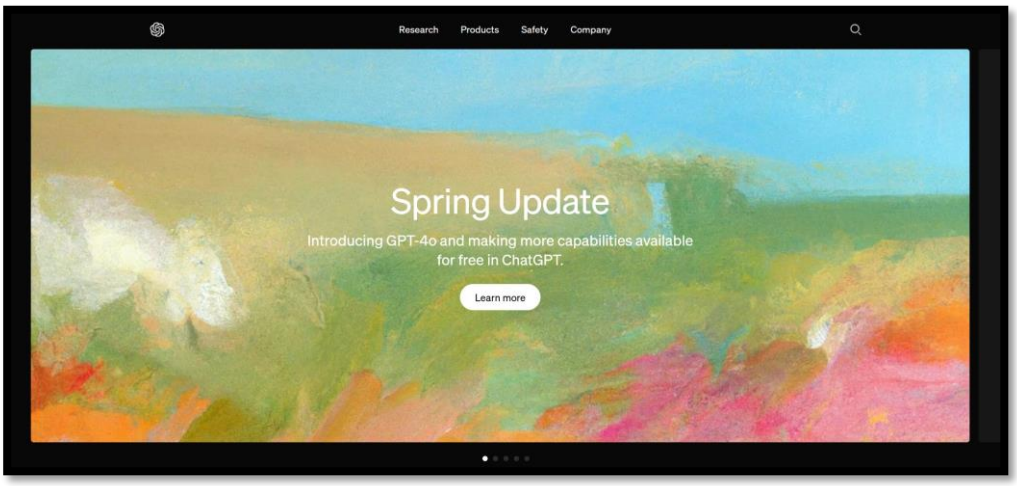
- **Encoder:** Captura a representação de contexto de cada palavra na sequência de entradaconsiderando toda a sequência.
- **Decoder:** Gera a sequência de saída palavra por palavra, utilizando a saída do encoder e oque foi gerado até o momento.

O sucesso dos Transformers foi o ponto de partida para o desenvolvimento de modelos em grande escala de linguagem, como BERT (Representações de Codificador Bidirecional de Transformers), GPT (Transformadores Geradores Pré-treinados), entre outros, que trouxeram novos benchmarks de desempenho para várias tarefas de PLN. Tais modelos pré-treinados podem ser afinados para qualquer tarefa em questão; portanto, uma tarefa complexa de PLN pode ser facilmente realizada com muito menos modelagem para essa tarefa.

5.2 OPENAI

OpenAI é um laboratório de pesquisa em IA baseado nos EUA, conhecido por métodos inovadores e ousados no campo. Fundada em dezembro de 2015, a OpenAI mostrou ter capacidades no desenvolvimento de tecnologia para empurrar as fronteiras da IA enquanto garante que a IA superinteligente resultante seja segura e benéfica para a humanidade. Na Figura 11 está a página inicial do site da empresa.

Figura 11 – OPENAI



Fonte: OPENAI, 2024

Um momento crucial no desenvolvimento da inteligência artificial foi a fundação da OpenAI. A empresa foi fundada com uma visão clara e um objetivo ambicioso por um grupo de tecnólogos e empreendedores proeminentes, incluindo Elon Musk, Greg Brockman e Ilya Sutskever. Como resultado da adesão inicial ao conselho de administração fornecida por indivíduos notáveis como Sam Altman, a OpenAI tornou-se uma força dominante na indústria de IA muito rapidamente.

O objetivo da OpenAI é promover a inteligência artificial geral (IAG) de uma forma que seja segura e vantajosa para todos. A IAG, que é definida como sistemas altamente autónomos que funcionam melhor que os humanos na maioria das tarefas economicamente relevantes, é um objetivo de longo prazo para a OpenAI, que visa equilibrar os avanços científicos com preocupações éticas e de segurança.

A OpenAI formou alianças e investimentos notáveis, incluindo um investimento notável de 1 bilhão de dólares da Microsoft em 2019. Com um investimento de 10 bilhoes de dólares em 2023, esta parceria estratégica foi alargada, indicando a confiança e a dedicação da Microsoft aos objectivos da OpenAI.

Entre os produtos criados pela OpenAI estão os modelos de linguagem GPT (Generative Pre-trained Transformer) GPT-3, 3.5 e 4, bem como o OpenAI Five, um esquadrão de IA para o jogo Dota 2. Além disso, novos benchmarks para capacidades de IA foram definidas por inovações como OpenAI Codex, que

auxilia na escrita de código, e DALL·E, que gera recursos visuais a partir de descrições textuais. Um modelo particularmente digno de nota é o ChatGPT, que simula muito bem as interações humanas, gerando texto de forma fluida e contextual.

A influência da empresa na comunidade de IA é inegável. Além de ultrapassar os limites tecnológicos, o seu trabalho e bens iniciam conversas significativas sobre o papel que a inteligência artificial desempenhará na sociedade no futuro. No futuro, a empresa provavelmente estará na vanguarda do desenvolvimento da tecnologia de IA, mesmo diante dos dilemas operacionais e morais que acompanham os seus avanços.

Em resumo, OpenAI é um excelente exemplo das possibilidades notáveis, bem como das dificuldades morais e práticas associadas à inteligência artificial contemporânea. A empresa deve garantir que os avanços que faz na tecnologia de IA sejam seguros e vantajosos para todas as pessoas, ao mesmo tempo que inova e molda o futuro do campo. A viagem da OpenAI, repleta de invenções, alianças empresariais e enigmas morais, será sem dúvida importante para determinar como a IA se desenvolverá no futuro.

5.3 GPT-3

GPT-3 é um modelo de inteligência artificial focado em compreender e gerar texto de maneira natural. Ele faz parte da terceira geração da família de modelos GPT, que usa a arquitetura de rede neural Transformer. Isso significa que ele foi projetado para entender o contexto de um texto e gerar respostas ou conteúdo novo que se alinhe a esse contexto.

A OpenAI, começou a desenvolver modelos GPT com o objetivo de avançar no campo do processamento de linguagem natural. Cada nova versão do GPT foi projetada para ser mais poderosa que a anterior, com o GPT-3 sendo notavelmente maior e mais sofisticado do que o GPT-2. O GPT-3 foi oficialmente lançado em junho de 2020 e representou um salto significativo em termos de capacidade de gerar texto coerente e contextualmente relevante.

O propósito principal do GPT-3, incluindo a versão 3.0, era avançar as fronteiras do que é possível com o processamento de linguagem natural. Isso

inclui:

- **Geração de Texto:** Produzir texto em diversos formatos e estilos, desde artigos e resumosaté poesia e código de programação.
- **Compreensão de Contexto:** Entender e responder a perguntas com base em um grandecorpus de conhecimento pré-treinado.
- **Tradução de Linguagem:** Traduzir texto entre idiomas de forma mais eficaz.
- **Assistência em Tarefas:** Auxiliar em tarefas que requerem compreensão ou geração delinguagem, como redação de e-mails, criação de conteúdo e programação.

O GPT-3.0, em particular, serviu como um marco demonstrando a viabilidade de modelos de linguagem em larga escala para uma ampla gama de aplicações, impulsionando o desenvolvimento de novos serviços e ferramentas baseadas em IA.

Apesar do fato de o GPT-3.0 ter avançado significativamente no campo do processamento de linguagem natural, ainda existem muitos obstáculos que devem ser superados antes que ele possa ser efetivamente utilizado. A existência de tendencias nos grandes conjuntos de dados de treinamento está entre os problemas mais importantes. A forma como o GPT-3.0 se comporta cria dilemas morais significativos sobre sua aplicação em ambientes práticos. A capacidade do modelo de gerar informações que parecem precisas, mas na verdade são imprecisas ou enganosas, apresenta outra dificuldade. Isto se deve ao fato de que o GPT-3.0, apesar de seus avanços, ainda não consegue confirmar de forma independente a integridade dos dados que aprendeu durante o treinamento. Isto pode apresentar desafios em aplicações onde são necessárias alta precisão e confiabilidade.

Outro desafio considerável do GPT-3.0 é sua tendência de gerar conteúdo sensível ou totalmente inapropriado, incluindo linguagem ofensiva ou até discursos de ódio. Isso se torna ainda mais complicado quando pensamos em usá-lo em ambientes que exigem um cuidado especial com o tipo de conteúdo produzido, como em contextos educacionais ou corporativos. Além dessas questões relacionadas ao conteúdo, enfrentamos também problemas práticos com o

GPT-3.0, como o alto consumo de recursos computacionais. Esses fatores não só encarecem seu uso, como também levantam questões importantes sobre sua viabilidade a longo prazo e impacto ambiental, o que nos faz questionar até que ponto podemos ou devemos escalar essa tecnologia.

5.4 GPT-3.5

O ChatGPT 3.5, uma iteração iterativa do modelo de linguagem GPT-3. “Generative Pre-trained Transformer” denota que o modelo é pré-treinado em uma extensa composição de dados textuais disponíveis na Internet e é baseado na arquitetura Transformer, especializada na produção de texto de forma autônoma. O número “3,5” alude a uma pequena melhoria ou atualização em relação à versão GPT-3.

Versões iterativas do GPT-3, como a 3.5, geralmente buscam aprimorar a capacidade do modelo de compreender e produzir linguagem natural de forma mais precisa e contextualizada. Isto aumenta a eficiência da utilização dos recursos informáticos, diminuir a geração de respostas tendenciosas ou inadequadas e melhorar a capacidade de compreensão de variantes linguísticas. Características Principais:

- **Treinamento em Larga Escala:** Semelhante aos seus antecessores, o GPT-3.5 pode produzir resultados em uma infinidade de campos do conhecimento, uma vez que é treinado em um vasto corpus de texto que cobre uma ampla gama de informações acessíveis online.
- **Capacidade de Generalização:** Graças ao seu extenso treinamento, o GPT-3.5 pode realizar uma ampla variedade de tarefas de processamento de linguagem natural (PLN) sem a necessidade de treinamento específico para tarefas, incluindo tradução de idiomas, geração de texto criativo, resumo de texto e até mesmo programação.
- **Melhorias Incrementais:** Embora detalhes específicos possam variar, é típico que atualizações como o GPT-3.5 introduzam melhorias no processamento de linguagem, precisão das respostas, entendimento contextual e capacidade de manter conversas coerentes e contextuais por períodos mais longos.

As aplicações para GPT-3.5 são muitas e incluem tudo, desde sofisticados sistemas de suporte à decisão e interfaces de programação automatizadas até assistentes virtuais inteligentes e ferramentas de produção de conteúdo. O GPT-3.5 tem sido empregado em diversas áreas, incluindo desenvolvimento de software, saúde, educação e entretenimento, devido à sua versatilidade e capacidade computacional.

Mesmo com seus recursos sofisticados, o GPT-3.5, como outros modelos de linguagem de grande escala, tem problemas de parcialidade, precisão de informações e produção de texto impróprio ou sensível assim como explicado no GPT-3.0. Para resolver estes problemas e aumentar a segurança e a utilidade dos modelos de linguagem, a OpenAI e a comunidade de IA ainda estão a pesquisar soluções.

5.5 GPT-4

ChatGPT-4.0 é uma evolução significativa do modelo de linguagem GPT-3, também desenvolvido pela OpenAI. Lançado em Janeiro de 2023, o ChatGPT-4.0 representa um avanço na série de modelos *Generative Pre-trained Transformer*, incorporando melhorias substanciais em relação às versões anteriores, tanto em termos de capacidade de processamento linguístico quanto de generalização de tarefas.

Tabela 1 – Comparação versões do GPT

Versão	**Modelo de Linguagem**	**Precisão**	**Abrangência**	**Criatividade**
ChatGPT 3.0	GPT-3.5	Boa	Razóavel	Boa
ChatGPT 3.5	GPT-4.0	Muito Boa	Muito Boa	Boa
ChatGPT 4.0	GPT-5.0	Excelente	Excelente	Excelente

Fonte: Autoria própria

- **Arquitetura Aperfeiçoada:** O design do *Transformer* ainda é usado pelo ChatGPT- 4.0, mas a arquitetura da rede, a contagem de parâmetros e as técnicas de treinamento foram significativamente melhoradas. Comparado ao GPT-3, que possui 175 bilhões de parâmetros, este modelo possui aproximadamente o dobro da capacidade. Produção e interpretação de texto mais avançadas são possíveis graças a esse aumento de parâmetros.
- **Treinamento Multimodal:** E a capacidade do ChatGPT-4.0 de lidar com pistas visuais em algumas situações além do texto o que o torna um modelo "multimodal", sendo essa uma de suas novas características. Como resultado, agora pode ser usado para trabalhos que exigem compreensão de textos e imagens.
- **Maior Contextualização e Precisão:** O modelo é mais capaz de sustentar discussões coerentes e contextuais por longos períodos e fornecer respostas mais precisas e abrangentes, uma vez que foi treinado para compreender e criar respostas tendo em conta um contexto mais amplo.

O ChatGPT-4.0 é usado em uma ampla gama de aplicações que se beneficiam de processamento avançado de linguagem natural e geração de conteúdo, incluindo:

- **Assistentes virtuais:** Melhorando a comunicação entre chatbots e assistentes pessoais por meio de trocas mais conversacionais e esclarecedoras.
- **Educação:** Ajudar no desenvolvimento de materiais instrucionais interativos, aulas individualizadas e respostas a perguntas frequentes.
- **Indústria Criativa:** Auxiliando autores e produtores de conteúdo na geração de conceitos, autoria de textos e até na composição de roteiros de jogos ou filmes.
- **Análise de Dados e Resumos:** Automatizando o processo de análise de grandes quantidades de material e criação de resumos educacionais claros.
- **Suporte ao Cliente:** fornecendo respostas rápidas e precisas às consultas de clientes em diversos setores, aumentando a satisfação do usuário e a eficácia operacional.

5.5.1 Plugins

O ChatGPT 4-0 também trouxe consigo a capacidade de utiliza plugins e APIs. Os plugins permitem que o ChatGPT execute uma variedade de tarefas práticas, tais como:

- Acesso a Informações Atualizadas: Com plugins, o ChatGPT pode buscar informações em tempo real, como previsões do tempo, valor de ações, ou notícias atualizadas, que não estão disponíveis em seu conjunto de treinamento padrão.
- Integração com Ferramentas e Plataformas: Os plugins possibilitam a integração com plataformas como serviços de reserva, sistemas de gerenciamento de conteúdo, e ferramentas de produtividade, permitindo ao ChatGPT executar tarefas específicas como agendar reuniões, reservar voos, ou manipular dados em softwares empresariais.
- Personalização de Respostas: Com a capacidade de conectar-se a bases de dados personalizadas e APIs específicas do domínio, o ChatGPT pode fornecer respostas e serviços mais personalizados aos usuários, ajustando-se a contextos específicos de negócios ou preferências pessoais.
- Interatividade Aumentada: A adição de plugins aumenta a interatividade do ChatGPT, permitindo diálogos mais complexos e funcionais, onde o usuário pode realizar ações efetivas durante a conversa.

A implementação de plugins no ChatGPT requer um cuidado a mais, especialmente no que diz respeito à segurança e à privacidade. A OpenAI implementou diretrizes rigorosas para os desenvolvedores de plugins, garantindo que eles aderem a padrões de segurança para proteger os dados dos usuários e manter a integridade do sistema. Os plugins devem ser explicitamente ativados pelos usuários, e estes têm controle total sobre quais plugins estão ativos durante suas sessões de interação com o ChatGPT.

A inclusão de plugins no ChatGPT abre muitas possibilidades para aplicativos de IA. Não só melhora a utilidade do ChatGPT em cenários de uso prático, mas também estabelece um precedente para futuras integrações entre modelos de IA e o vasto ecossistema de serviços digitais e dados online. À

medida que mais desenvolvedores criam e integram plugins, espera-se que o ChatGPT continue evoluindo, tornando-se cada vez mais uma ferramenta versátil e indispensável em muitos aspectos da vida digital e real.

Sendo uma tecnologia muito poderosa, o ChatGPT-4.0 também levanta questões éticas de grande importância, como a privacidade dos dados conforme mencionado, a segurança e a preocupação com a capacidade de criar informações incorretas ou mesmo maliciosas. A OpenAI trabalhou para implementar e incentivar ativamente o uso ético da tecnologia para ajudar a mitigar as consequências dessas preocupações. Em termos mais amplos, o ChatGPT-4.0 é um exemplo do impacto em curso da IA na natureza da interação com nossas tecnologias, trazendo ferramentas à vida que são capazes de se envolver de maneiras cada vez mais dinâmicas e úteis.

5.6 GPT-4 Turbo

Os modelos de linguagem tornaram-se mais sofisticados e potentes devido aos avanços contínuos na tecnologia de inteligência artificial (IA). lançado em Novembro de 2023 um modelo notável é o modelo GPT-4T desenvolvido pela OpenAI, que é uma variante personalizada do modelo GPT-4. Na busca por uma IA totalmente adaptável e prática, este modelo representa um grande avanço, uma vez que foi criado para lidar com determinados trabalhos que exigem um grau de precisão e compreensão além do que os modelos gerais podem fornecer.

Em resposta à crescente necessidade de aplicações mais precisas e direcionadas em todo o amplo domínio da inteligência artificial, foi criado o GPT-4T. O GPT-4T é derivado da plataforma GPT-4 estável da OpenAI e tem como objetivo principal melhorar a compreensão técnica e as habilidades de tradução do modelo básico. Um grupo de indivíduos proeminentes, incluindo Ilya Sutskever e outros pesquisadores importantes da OpenAI, liderou esse progresso.

Embora o GPT-4T seja semelhante ao seu antecessor em termos de design do transformador, ele difere porque fez algumas modificações na configuração da camada e na metodologia de treinamento, permitindo melhor desempenho em domínios como interpretação científica e tradução técnica. O principal avanço tecnológico do GPT-4T é a sua precisão superior na compreensão e interpretação

da linguagem técnica em comparação com versões anteriores.

A OpenAI usou extensos materiais textuais especializados com algoritmos de aprendizado de máquina de última geração para treinamento GPT-4T. Durante a fase de aprendizagem, o modelo foi exposto a um grande número de publicações técnicas e científicas, com ênfase em dados cuidadosamente selecionados para minimizar vieses e aumentar a precisão.

O GPT-4T tem sido usado em vários campos onde é necessária grande precisão linguística. É útil na medicina, por exemplo, ao escrever diagnósticos e compreender documentos médicos. Auxilia na tradução de manuais técnicos e documentos de patentes na área de engenharia.

Possuindo capacidades sofisticadas, mas também precisa ser atualizado com frequência para permanecer bem-sucedido devido à sua complexidade operacional. Além disso, a utilização de dados sensíveis para formação em IA levanta questões éticas e de privacidade.

Sendo também uma excelente ilustração de como a especialização do modelo de IA pode satisfazer demandas específicas de maneira surpreendente e eficaz. Serve como exemplo do avanço da tecnologia de IA e enfatiza o valor de abordagens especializadas para questões desafiadoras.

5.7 GPT-4 omni

Um passo em direção a uma interação humano-computador muito mais natural, o GPT-4o ("o" de "omni") aceita qualquer combinação de texto, voz, imagem e vídeo como entrada e produz qualquer combinação de saídas de texto, áudio e imagem. Seu tempo de reação às entradas de áudio é em média de 232 milissegundos, com máximo de 320 milissegundos; isso é comparável ao tempo de resposta humano típico (abre em uma nova janela) durante uma conversa. Além de ser muito mais rápido e mais barato na API, ele se iguala ao desempenho do GPT-4 Turbo em texto e código em inglês e melhora significativamente o texto em idiomas diferentes do inglês. Comparado com outras versões, o GPT-4o se destaca na compreensão visual e auditiva.

O modo de voz permitiu conforme Figura 12 que você se comunicasse com ChatGPT com latências médias de 2,8 segundos (GPT-3,5) e 5,4 segundos (GPT-

4) antes do GPT-4o. O Modo de Voz faz isso utilizando um pipeline que consiste em três modelos distintos: um modelo básico que converte áudio em texto, um modelo GPT-3.5 ou GPT-4 que recebe entrada de texto e o envia e um terceiro modelo básico que transforma o texto de volta. em áudio. Como resultado deste procedimento, a fonte primária de inteligência, GPT-4, perde muitas informações. Ele é incapaz de perceber o tom, vários alto-falantes, ruído de fundo ou emoções como rir ou cantar.

Figura 12 – chatgpt-4 omni

Fonte: autória própria

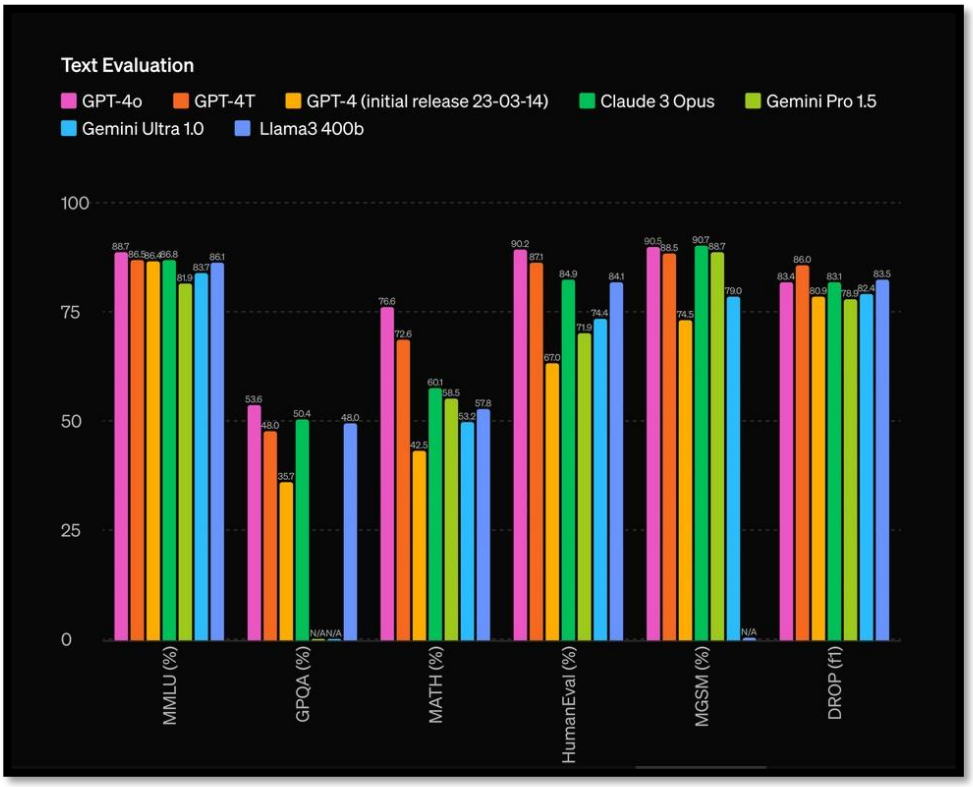
Ao usar o GPT-4o, a OPENAI conseguiu treinar um único novo modelo de ponta a ponta para texto, visão e áudio, o que significa que a mesma rede neural lida com todas as entradas e saídas. Como o GPT-4o é o primeiro modelo da empresa a incorporar todas essas modalidades, estão apenas começamos a explorar as capacidades e restrições do modelo.

Conforme medido em benchmarks tradicionais, o GPT-4o alcança desempenho de nível GPT-4 Turbo em texto, raciocínio e inteligência de codificação, ao mesmo tempo que estabelece novos padrões elevados em recursos multilíngues, de áudio e de visão.

- **MMLU (*Massive Multitask Language Understanding*)**: Esse benchmark foi projetado para medir a capacidade do modelo de compreender e responder a um conjunto diversificado de questões de natureza educacional e profissional distribuídas em muitos domínios de conhecimento.
- **GPQA (*General Purpose Question Answering*)**: Esse benchmark foi projetado para avaliar o nível geral e a amplitude do conhecimento dos modelos acima de uma grande diversidade de áreas e tipos de perguntas, para verificar sua utilidade em aplicações do mundo real, onde as perguntas podem ser amplamente variadas e bastante inesperadas.
- **MATH**: Benchmark projetado para avaliar a capacidade dos modelos de resolver problemas matemáticos usando apenas texto, sem a necessidade de cálculos numéricos externos, não pôde ser processado devido ao limite mínimo de palavras. Por favor, forneça um texto mais longo para que eu possa ajudar a parafrasear corretamente.
- **HumanEval**: Esse benchmark testa a capacidade dos modelos de programação de gerar código correto e eficiente para uma variedade de problemas de programação. É uma forma de medir a habilidade de um modelo em entender e completar tarefas de codificação.
- **MGSM (Multi-Genre Semantic Matching)**: Mede a capacidade de um modelo de linguagem de entender e relacionar informações entre diferentes gêneros ou tipos de textos.
- **DROP (Discrete Reasoning Over the content of Paragraphs):** É um benchmark que desafia os modelos a realizarem raciocínio discreto e a responder a perguntas que requerem manipulação de várias partes de informações contidas em um ou mais parágrafos.

Na Figura 13 podemos ver o comparativo do GPT-4º em relação a suas outras versões e ferramentas semelhantes.

Figura 13 – GPT-4o comparações



Fonte: OPENAI, 2024

Aprimoramento do raciocínio: GPT-4o atinge um novo pico de 88,7 por cento no COT MMLU de 0 disparo (questões de conhecimentos gerais). Nova biblioteca foi usada para coletar todas essas avaliações. Além disso, o GPT-4o atinge uma nova pontuação alta de 87,2 por cento no MMLU convencional de 5 disparos sem CoT. Llama3 400b está atualmente em treinamento.

Tabela 2 – Comparação versões do GPT-4

Versão	**Características**	**Pontos Fortes**	**Pontos Fracos**
GPT-4	Alto Desempenho e raciocínio	Precisão e fluidez	Tempo de resposta alto, custo computacional alto
GPT-4t	Velocidade e eficiência	Rapidez na gera ção de texto	Menor precisão e capacidade de raciocínio

GPT-4o	Otimização tarefas espe cificas	Melhor desempenho em áreas específicas	Desempenho inferior em tarefas gerais

Fonte: Autoria própria

6 CONSIDERACÕES FINAIS

Para responder a seguinte questão de pesquisa: como a evolução da tecnologia tornou possível treinar e executar modelos como o ChatGPT?

O percurso impressionante da inteligência artificial, desde suas origens com projetos pioneiros como a ELIZA até os avanços contemporâneos representados pelo ChatGPT. Este trajeto não apenas demonstra o progresso técnico na área, mas também demonstra a crescente integração da IA no dia-a-dia, redefinindo interações e moldando novas realidades digitais.

O desenvolvimento contínuo de sistemas de aprendizado de máquina, especialmente através de técnicas como redes neurais e processamento profundo de linguagem natural, promete avanços ainda mais significativos. Estes sistemas são cada vez mais capazes de realizar tarefas complexas que antes eram exclusivas aos seres humanos, desde a condução autônoma até a interação sofisticada por meio de chatbots.

Reconhecer também como o aprendizado de máquina serve como um pilar fundamental para esses avanços, fornecendo as ferramentas e métodos que permitem que a IA não apenas imite, mas também expanda a capacidade humana de resolver problemas e entender o mundo ao nosso redor. Este reconhecimento prepara o cenário para uma exploração mais profunda das capacidades e desafios do aprendizado de máquina, destacando sua importância crítica no desenvolvimento contínuo da inteligência artificial.

Nesse sentido, o ChatGPT é um dos grandes exemplos de capacidade na IA até agora. A capacidade quase humana de compreender e gerar linguagem não é apenas prova direta do quão avançados tecnologicamente estamos, mas também do que vem a seguir. Oferecendo uma prévia do tipo de interação que poderíamos ter com as máquinas no futuro: imaginar um assistente digital que não só compreenda o seu significado, mas adivinhe as suas necessidades, mesmo antes de você percebê-las como a nova versão do chatgpt-4 omni vem realizando.

Inovações como o ChatGPT estão empurrando as fronteiras do que é possível. Desafiam a compreensão e as capacidades de uma máquina e estabelecem uma nova fronteira sobre o que IA será capaz no futuro.

Portanto, como trabalho futuro, este estudo tem como objetivo analisar e avaliar o hardware afundo para um conhecimento mais avancado referente ao estudo das inteligência artificial como o ChatGPT.

REFERÊNCIAS

BAROCAS, S.; SELBST, A. D. Big Data's Disparate Impact. **California Law Review**, 2016.

BENGIO, Y. **Deep Learning of Representations: Looking Forward. Statistical Language and Speech Processing**. [s.l: s.n.].

BENGIO, Y.; LECUN, Y.; HINTON, G. Deep learning for AI. **Communications of the ACM**, v. 64, n. 7, p. 58–65, 2021.

BOSTROM, N.; YUDKOWSKY, E. **The Ethics of Artificial Intelligence. Cambridge Handbook of Artificial Intelligence**. [s.l: s.n.].

BOYD, D.; CRAWFORD, K. **Critical Questions for Big Data. Information**. [s.l.] Communication & Society, 2012.

BROWN, T. B. et al. **Language models are few-shot learners**. [s.l: s.n.].

BURRELL, J. How the machine 'thinks': Understanding opacity in machine learning algorithms. **Big Data & Society**, 2016.

Deep Learning. Disponível em: <https://www.deeplearningbook.org/>. Acesso em: 22 set. 2023.

EYSENCK, M. W.; EYSENCK, C. **Inteligência artificial × humanos: o que a ciência cognitiva nos ensina ao colocar frente a frente a mente humana e a IA [recurso eletrônico]. Tradução: Gisele Klein. Revisão técnica: Vitor Geraldi Haase**. Porto Alegre: Artmed, 2023.

FLORIDI, L. et al. AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines. v. 28, p. 689–707, 2018.

FØLSTAD, A.; BRANDTZÆG, P. B. Chatbots and the new world of HCI. **interactions**, v. 24, n. 4, p. 38–42, 2017.

GABRIEL, M. **Inteligência artificial: do zero ao metaverso**. Barueri [SP]: Atlas, 2024.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H. **The elements of statistical learning: Data mining, inference, and prediction**. 2. ed. Nova Iorque, NY, USA: Springer, 2009.

HAYKIN, S. S. **Neural Networks and Learning Machines. 3**. Upper Saddle River, NJ: Pearson Education, 2009.

HINTON, G. et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. **IEEE Signal Processing Magazine**, v. 29, n. 6, p. 82–97, 2012.

HINTON, G. E.; OSINDERO, S.; TEH, Y. W. A Fast Learning Algorithm for Deep Belief Nets. **Neural Computation**, 2006.

Introducing GPT-4o: our fastest and most affordable flagship model. Disponível em: <https://platform.openai.com/docs/guides/chat>. Acesso em: 15 abr. 2024.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. **Science (New York, N.Y.)**, v. 349, n. 6245, p. 255–260, 2015.

KAPLAN, A. From Chatbots to ChatGPT: The Evolution of Conversational AI. **Journal of Artificial Intelligence Research**, 2021.

KRIZHEVSKY, A. ImageNet Classification with Deep Convolutional Neural Networks. NIPS. p. 1106–1106, 2012.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. **Communications of the ACM**, v. 60, n. 6, p. 84–90, 2017.

LAU, J. H. Chatbots to ChatGPT in a Cybersecurity Space: Evolution and Challenges. **Journal of Cybersecurity**, 2020.

LUGER, G. F. **Artificial intelligence: Structures and strategies for complex problem solving: International edition**. Upper Saddle River, NJ, USA: Pearson, 2009.

LUND, B. et al. ChatGPT and a new academic reality: AI-written research papers and the ethics of the large language models in scholarly publishing. **SSRN Electronic Journal**, 2023.

PAPINENI, K. et al. **BLEU: A method for automatic evaluation of machine translation**. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02. **Anais**...Morristown, NJ, USA: Association for Computational Linguistics, 2001.

PENG, H. et al. **Evaluating Emerging AI/ML Accelerators: IPU, RDU, and NVIDIA/AMD GPUs**. Companion of the 15th ACM/SPEC International Conference on Performance Engineering. **Anais**...New York, NY, USA: ACM, 2024.

RADFORD, A. et al. **Language Models are Unsupervised Multitask Learners**. Disponível em: <https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf>. Acesso em: 1 mar. 2024.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, v. 65, n. 6, p. 386–408, 1958.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: A modern approach**. [s.l.] Prentice Hall, 2010.

Unidades de procesamiento de tensor (TPUs). Disponível em: <https://cloud.google.com/tpu>. Acesso em: 5 nov. 2023.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS
GABINETE DO REITOR

Av. Universitária, 1069 • Setor Universitário
Caixa Postal 86 • CEP 74605-010
Goiânia • Goiás • Brasil
Fone: (62) 3946.1000
www.pucgoias.edu.br • reitoria@pucgoias.edu.br

RESOLUÇÃO n° 038/2020 – CEPE

ANEXO I

APÊNDICE ao TCC

Termo de autorização de publicação de produção acadêmica

O(A) estudante Matheus Afonso B. da Silva
do Curso de Ciências da Computação, matrícula 2016.1.0008.0262-4,
telefone: 62 9.8178-0899 e-mail Matheusjogosdb@outlook.com
na qualidade de titular dos direitos autorais, em consonância com a Lei nº 9.610/98 (Lei dos Direitos do Autor), autoriza a Pontifícia Universidade Católica de Goiás (PUC Goiás) a disponibilizar o Trabalho de Conclusão de Curso intitulado
D. Elza ao ChatGPT: História e Evolução da I.A.
_____, gratuitamente, sem ressarcimento dos direitos autorais, por 5 (cinco) anos, conforme permissões do documento, em meio eletrônico, na rede mundial de computadores, no formato especificado (Texto(PDF); Imagem (GIF ou JPEG); Som (WAVE, MPEG, AIFF, SND); Vídeo (MPEG, MWV, AVI, QT); outros, específicos da área; para fins de leitura e/ou impressão pela internet, a título de divulgação da produção científica gerada nos cursos de graduação da PUC Goiás.

Goiânia, 26 de Maio de 2024.

Assinatura do autor: Matheus Afonso B. da Silva

Nome completo do autor: Matheus Afonso B. da Silva

Assinatura do professor–orientador: _____

Nome completo do professor-orientador: _____