



# Generative Learning e LLM

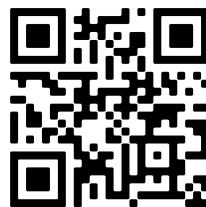
Eldrey S. Galindo



# Quem sou eu?

---

- Especialista em Deep Learning - DeepLearning.AI
  - Mestre Ciências da Computação - UNICAMP
  - Bacharel em Sistemas da Informação - UTFPR
- 
- Technical Data Scientists Manager - CESAR (9 anos)
  - Professor da graduação de Ciências da Computação - CESAR School (2 anos)



Página Pessoal





C . E . S . A . R



- **+25 anos**  
experiência em inovação, tecnologia, design e empreendedorismo, com desenvolvimento de soluções robustas e atuando no ciclo completo da inovação
- **+1150**  
técnicos e +100 consultores
- **+200**  
startups apoiadas em diversos modelos
- **+140 projetos**  
Executados por ano, em + 20 segmentos
- **CESAR School**  
cursos de graduação, especialização, mestrado e doutorado em ciências da computação, eng. de software e design

# Visão geral

## Artificial Intelligence, AI

(Generic term, mimic cognitive functions)

## Machine Learning, ML

(Number of samples > 100)

Linear Regression, Random Forest, Decision Trees, Gaussian Processes, ...

## Deep Learning, DL

(Number of samples > 500)

Artificial Neural Network (ANN), Convolution Neural Network (CNN), Graph Neural Network (GNN), Variational Encoders (VAE), Generative Adversarial Network (GAN), Recurrent Neural Network (RNN), Deep Reinforcement Learning (DRL), ...

Chemical Formula,  
SMILES, Fragments

Atomic Structure  
(Molecules, Solids,  
Proteins)

Text/Literature

XRD, XAS, Raman, NMR,  
UV-vis, XANES,  
Electron/Phonon DOS

SEM, STM, STEM images



*"Na natureza nada se cria, nada se  
perde, tudo se transforma"*

(Antoine Lavoisier)

# Visão geral

---

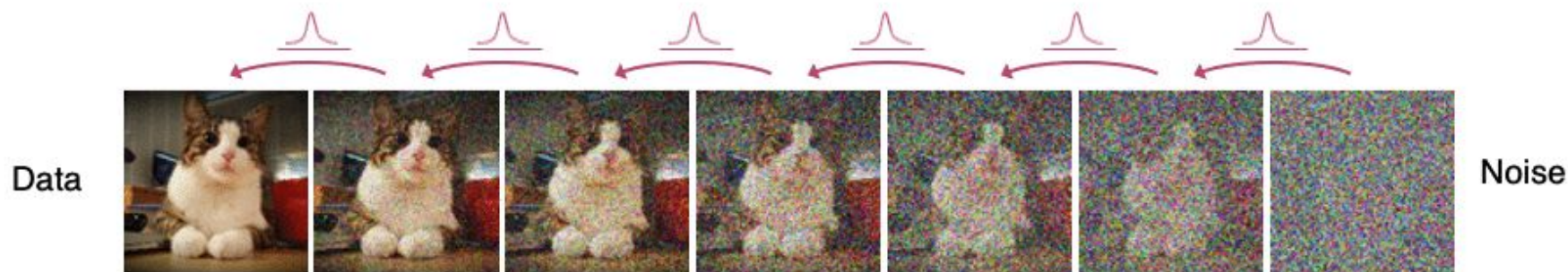


Generative learning é uma abordagem de aprendizado de máquina que envolve a criação de modelos que podem gerar novos dados com base em um conjunto existente de dados de treinamento. Em outras palavras, o modelo é capaz de aprender as características e padrões do conjunto de dados de treinamento e, em seguida, criar novos exemplos que se assemelham a esses dados originais. Isso pode ser útil em várias aplicações, como na geração de texto, imagens, vídeos e até mesmo em áreas como a medicina para a criação de novos medicamentos.





## ➡ Denoising Process with Unimodal Gaussian Distribution



## Denoising Process with Our Multimodal Conditional GAN



# Large Language Model (LLM)

Modelos que tentam prever a próxima palavra em um texto com base em um contexto

- Básico:
  - Autocompletar do Word, Android, "Você quis dizer..." Buscadores.
- Instrução:
  - Dado um contexto e instruções gere algo com base no contexto





# Bard AI

<https://bard.google.com/?hl=en>



<https://ai.facebook.com/blog/large-language-model-llama-meta-ai/>



# ChatGPT

## List of Open Sourced Fine-Tuned Large Language Models (LLM)

An incomplete list of open-sourced fine-tuned Large Language Models (LLM) you can run locally on your computer



Sung Kim · Follow

Published in Geek Culture · 21 min read · Mar 30

<https://medium.com/geekculture/list-of-open-sourced-fine-tuned-large-language-models-llm-8d95a2e0dc76>

# Chat GPT web



- É uma aplicação, feita para ser usada e para gerar engajamento
- A aplicação utiliza 1 versão do modelo, atualmente (01/05/23) GPT-3.5
- "Gratuito"
- As conversas ~~vão~~ podem ser usadas para re-treinamentos e melhorias
- Filtros, muitos filtros de respostas

# Chat GPT API

- É uma API
- Tem custo, extremamente baixo, mas tem
- Os dados não devem ser utilizados para re-treinamentos, segundo a OpenAI, esses dados são armazenados temporariamente
- Você pode usar qualquer um dos modelos disponíveis (8 modelos em 01/05/23)

# Quais modelos?



<a href="#"><u>GPT-4</u></a> Limited beta	A set of models that improve on GPT-3.5 and can understand as well as generate natural language or code
<a href="#"><u>GPT-3.5</u></a>	A set of models that improve on GPT-3 and can understand as well as generate natural language or code
<a href="#"><u>DALL·E</u></a> Beta	A model that can generate and edit images given a natural language prompt
<a href="#"><u>Whisper</u></a> Beta	A model that can convert audio into text
<a href="#"><u>Embeddings</u></a>	A set of models that can convert text into a numerical form
<a href="#"><u>Moderation</u></a>	A fine-tuned model that can detect whether text may be sensitive or unsafe
<a href="#"><u>GPT-3</u></a>	A set of models that can understand and generate natural language
<a href="#"><u>Codex</u></a> Deprecated	A set of models that can understand and generate code, including translating natural language to code

# O que é possível fazer?



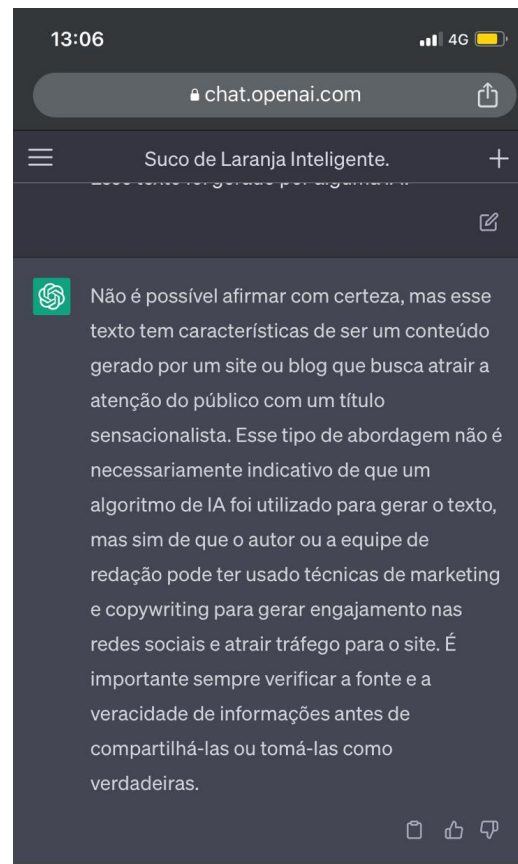
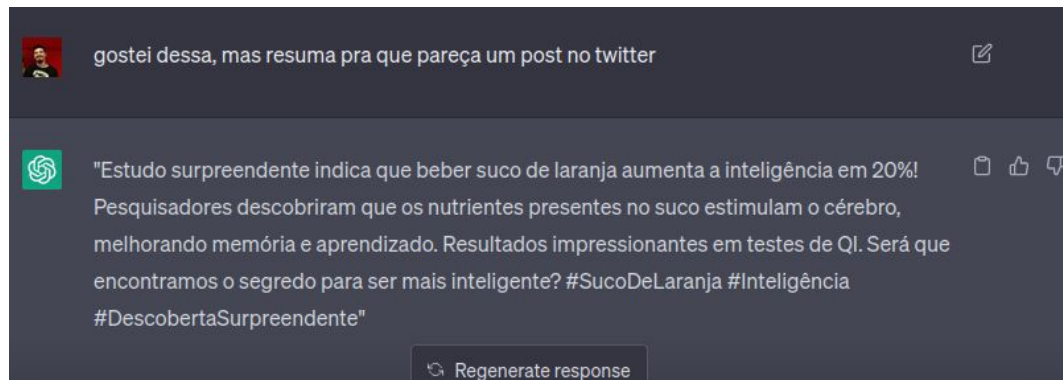
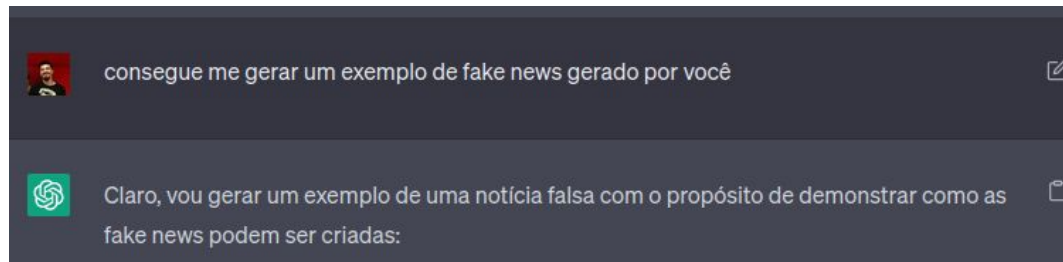
- Processamento de linguagem natural (NLP) e compreensão de texto.
  - Sumarização, detecção de sentimentos e contexto, etc
- Geração de texto, incluindo redação de artigos e criação de histórias.
  - Escrita de e-mail, completar textos (palavras, frases, parágrafos), geração de texto por estilo, etc
- Suporte ao cliente e chatbots inteligentes.
  - FAQ inteligente, Chatbot automático
- Tradução automática e interpretação de idiomas.
  - Tradução, correção de ortografia, ajuste de texto por estilo

# Ética, cuidados e fake news



"Estudo surpreendente indica que beber suco de laranja aumenta a inteligência em 20%! Pesquisadores descobriram que os nutrientes presentes no suco estimulam o cérebro, melhorando memória e aprendizado. Resultados impressionantes em testes de QI. Será que encontramos o segredo para ser mais inteligente? #SucoDeLaranja #Inteligência #DescobertaSurpreendente"

# Ética, cuidados e fake news





# Ética, cuidados e fake news



- Cuidado com dados sensíveis:
  - Se os LLMs forem re-treinados (fine tune) com dados sensíveis nada os impede de vazarem esses dados em uma resposta
  - Não envie dados sensíveis para nenhuma API de IA.
- Fake news:
  - Como um gerador de texto capaz de replicar estilos, ele consegue gerar qualquer notícia.
- Assinado IA:
  - Sempre que usar uma IA para gerar algo, diga que aquilo foi gerado por IA. (Seja ético)
- Viés:
  - Qualquer IA pode refletir os vieses dos dados, alguns podem ser preconceituosos ou discriminatórios



# Obrigado!

