

Fundamentos da Inteligência Artificial Generativa

O Dilema dos Dados:

Por que um Novo Jeito de Aprender?

● **Aprendizado Supervisionado Tradicional**

- Funciona como um aluno com um professor ao lado, explicando todas as regras e corrigindo os erros.
- Grande Desafio: Exige enormes volumes de dados rotulados por humanos, um processo caro e, muitas vezes, inviável.

● **Aprendizado Não Supervisionado**

- É como receber um jogo desconhecido sem manual ou instruções.
- O objetivo é encontrar padrões por conta própria, como agrupar peças semelhantes.





A Solução Elegante: Aprendizado Autossupervisionado (AAS)

● O que é?

- Uma abordagem "autodidata".
- O modelo cria seus próprios desafios e tarefas (chamadas "tarefas de pretexto") a partir dos dados, sem precisar de rótulos externos.

● Como funciona?

- A "supervisão" vem da própria estrutura dos dados.



O Coração dos Modelos Generativos: Entendendo a Realidade

● Função de Distribuição de Probabilidade (FDP)

- É a ferramenta matemática que descreve a chance de todos os resultados possíveis ocorrerem.
- Para um modelo de IA, aprender a FDP é como criar um "mapa da realidade" dos dados.

● O Objetivo do Modelo Generativo

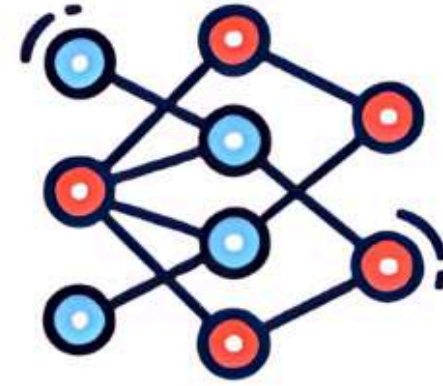
- Aprender a FDP dos dados de treinamento ($p(x)$) para poder gerar novas amostras realistas.

● Amostragem (Sampling)

- É o ato de "sortear" um novo dado que segue as regras da FDP que o modelo aprendeu.
- Exemplo: Se o modelo aprendeu a FDP de rostos, ele pode "amostrar" um $\text{imagem}_{\text{nova}}$.

Sinais de controle

Pontos de dados potenciais



Modelo generativo

Novos pontos de dados

Valores de probabilidade

Modelos Autorregressivos

Prevendo o Próximo Passo

● A Ideia Central

- Geram dados de forma sequencial, onde cada novo elemento é previsto com base em todos os elementos anteriores.

● A Tarefa Autossupervisionada

- A tarefa de pretexto é simples: prever o próximo elemento da sequência.
- O "rótulo" é o elemento que realmente vem a seguir nos dados de treinamento, sem precisar de anotação humana.



Aplicação Principal dos Modelos AR: Modelos de Linguagem (LLMs)

● O que são os LLMs (ChatGPT, Gemini, etc.)

- Em sua essência, são modelos de linguagem autorregressivos gigantesco.
- São treinados em quantidades massivas de texto usando a arquitetura Transformer

● Como eles Geram Texto?

1. Começam com uma instrução inicial (o 2. prompt).
3. Calculam a distribuição de probabilidade para o próximo token (palavra).
4. Selecionam um token e o anexam à sequência.
5. Repetem o processo, usando a nova sequência como base para prever o próximo token.



Outras Aplicações e os Prós e Contras

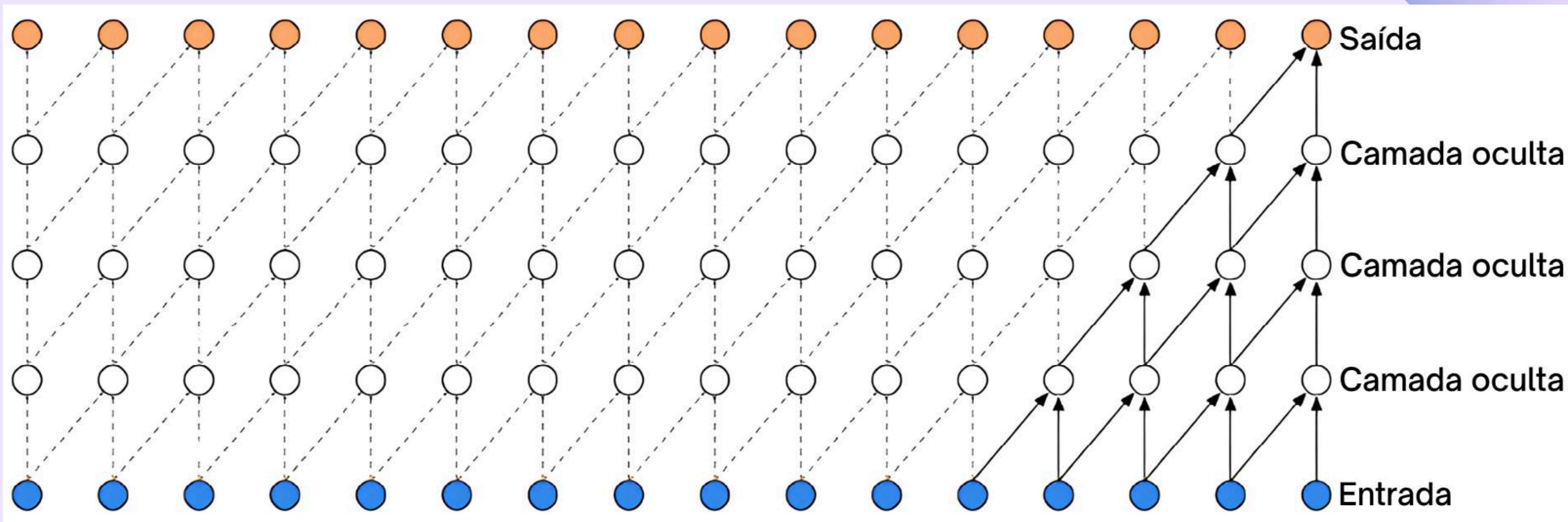
● Outras Aplicações

- Geração de Áudio (WaveNet): Preveem a próxima amostra de uma forma de onda de áudio, milhares de vezes por segundo.
- Séries Temporais: Previsão de valores futuros com base em dados históricos.
- Geração de Vídeo: Visto como uma sequência de quadros, onde se prevê o próximo quadro com base nos anteriores.

● Vantagens e Desvantagens

- Vantagem: Base probabilística clara; produz sequências muito coerentes.
- Desvantagem: A geração é inerentemente lenta e sequencial, dificultando a paralelização e sendo custosa para sequências longas.

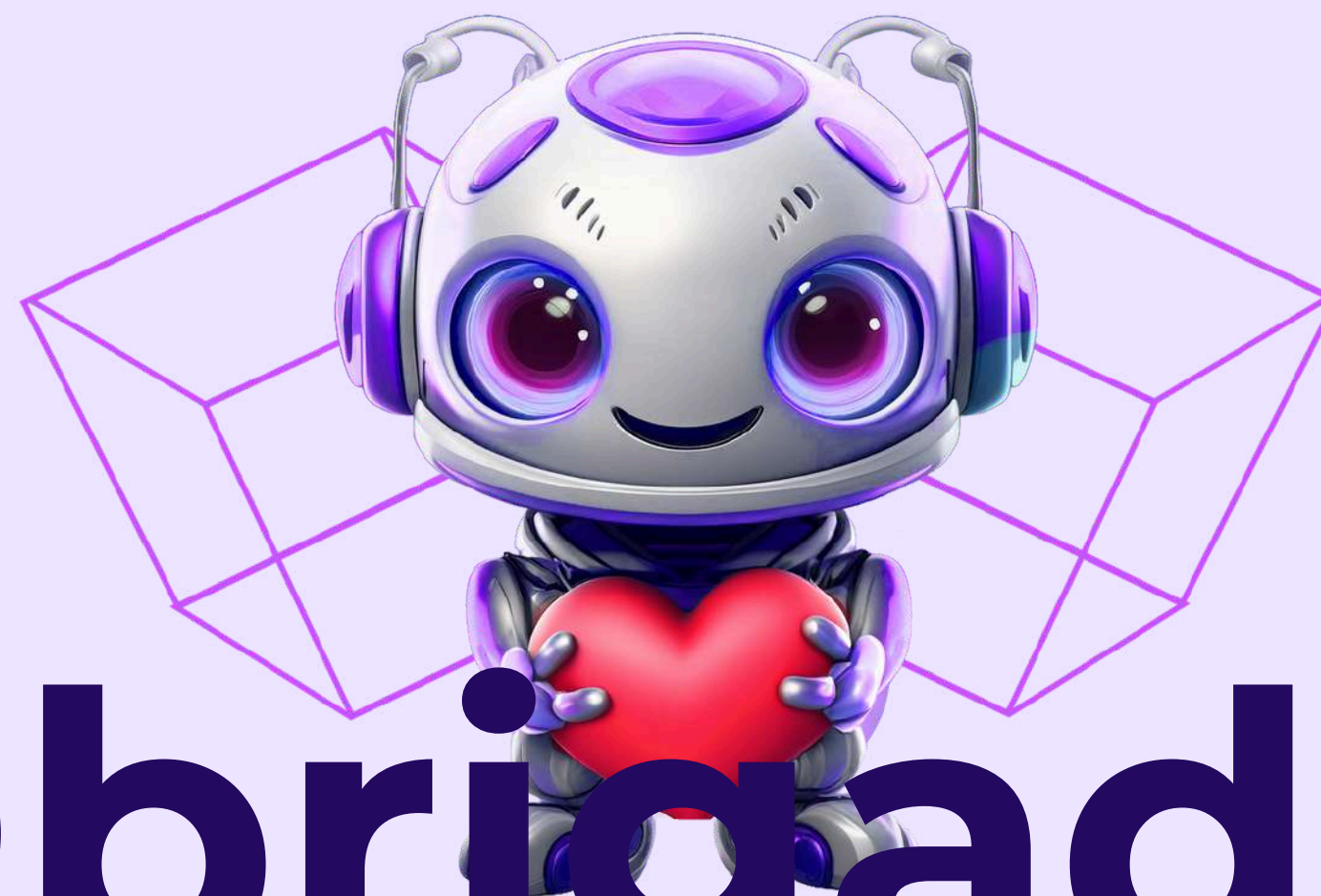




Recapitulação do que foi visto

- **Aprendizado Autossupervisionado:**
 - Nos permite treinar modelos sem a necessidade de rótulos humanos, usando os próprios dados como supervisão
- **Modelos Generativos:**
 - Aprendem a "distribuição de probabilidade" da realidade para criar novos dados.
- **Modelos Autorregressivos:**
 - São uma classe de modelos generativos que criam sequências passo a passo e são a base dos modernos LLMs.





Obrigada!

Perguntas?