



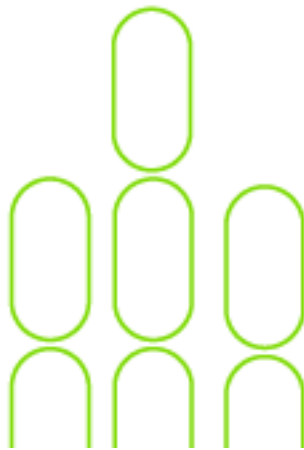
# Bootcamp: Cientista de Dados IA Expert

## Atividade Modular

Módulo 3: Avanços em AI Generativas e Suas Aplicações para a Ciência de Dados

### Objetivos de Ensino

Exercitar os seguintes conceitos trabalhados no Módulo:

1. **Entendimento e aplicação de Redes Generativas Adversariais (GANs):**  
Compreender a teoria e prática por trás das GANs, incluindo suas componentes principais – o gerador e o discriminador.
  2. **Manipulação e visualização de grandes datasets:** aprender a importar, processar e visualizar datasets complexos, preparando-os para uso em modelos de aprendizado de máquina.
  3. **Desenvolvimento e treinamento de modelos em ambientes controlados:** modelar e treinar GANs para gerar novas amostras de dados que imitam as características dos dados originais.
  4. **Avaliação crítica de modelos de IA:** aprender a avaliar a qualidade dos dados gerados por modelos de IA, usando métricas de performance específicas para GANs.
- 



## Enunciado

O projeto visa explorar o potencial das Redes Generativas Adversariais (GANs) para a geração de dados sintéticos baseados em classes específicas de imagens do dataset do [Google's Quick Draw](https://console.cloud.google.com/storage/browser/quickdraw_dataset/full/numpy_bitmap) ([https://console.cloud.google.com/storage/browser/quickdraw\\_dataset/full/numpy\\_bitmap](https://console.cloud.google.com/storage/browser/quickdraw_dataset/full/numpy_bitmap)).


O Quick Draw é um conjunto de dados composto por milhões de desenhos à mão, agrupados em centenas de categorias. Para este Trabalho Prático, você selecionará uma categoria específica e desenvolverá uma GAN para gerar novas imagens que se assemelhem aos desenhos originais dessa categoria. Para auxiliar no desenvolvimento da GAN, será disponibilizado um jupyter notebook (.ipynb) com o nome “trabalho\_pratico\_guia.ipynb”, assim como um arquivo com o nome dos pacotes necessários para rodar o notebook localmente.


## Tarefas

### 1. Seleção e Preparação do Dataset:

- Acesse o [Quick Draw Dataset](#).
- Escolha uma categoria de interesse (por exemplo, “gatos”, “árvores”, “carros” etc.) e baixe os dados correspondentes.
- Carregue e visualize uma amostra dos desenhos para entender as características gerais dos dados.

### 2. Desenvolvimento do Modelo GAN:

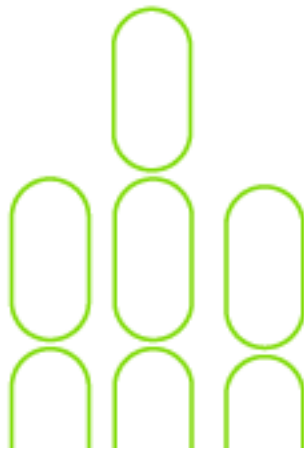
- **Modelagem do Gerador:** desenvolva um gerador que aprenda a criar imagens novas que imitem os desenhos da categoria escolhida. O gerador deve ser capaz de receber um vetor de ruído aleatório e transformá-lo em uma imagem.
- 

- 
- **Modelagem do Discriminador:** desenvolva um discriminador cuja função é distinguir entre os desenhos reais (do dataset) e as imagens geradas pelo gerador.
  - **Função de Perda:** implemente as funções de perda adequadas para treinar tanto o gerador quanto o discriminador, geralmente envolvendo a minimização do erro do discriminador em classificar as imagens corretamente e maximização do erro ao classificar imagens falsas como reais.

### 3. Treinamento do Modelo:

- **Treine sua GAN:** utilize os desenhos reais como base. Acompanhe a evolução do treinamento visualizando tanto as perdas do gerador quanto do discriminador, além de exemplos de imagens geradas durante o processo.

### 4. Avaliação e Análise:

- **Avalie o desempenho da sua GAN:** escreva como as imagens geradas se comparam às originais em termos de qualidade e diversidade.
- 



## Atividades

### **Atividade 1: Análise Exploratória Inicial e Preparação dos Dados**

**Descrição:** Acesse o [Quick Draw Dataset](#) e selecione uma categoria específica de interesse. Após o download dos dados, deve carregar e visualizar uma amostra significativa dos desenhos para entender as características gerais e a variação dentro da categoria escolhida. Esta atividade envolve a exploração inicial dos dados usando técnicas de visualização de dados e estatística descritiva para formar uma base sólida para o desenvolvimento do modelo.

### **Atividade 2: Construção e Configuração do Modelo GAN**

**Descrição:** Com base na compreensão adquirida na Atividade 1, construa o modelo GAN, que inclui tanto o gerador quanto o discriminador. O modelo pode ser desenvolvido utilizando como base a primeira aula interativa.

### **Atividade 3: Treinamento e Monitoramento do Modelo**

**Descrição:** Treine a GAN utilizando os dados reais como referência. Durante o treinamento, é crucial monitorar a evolução do desempenho tanto do gerador quanto do discriminador. Deverá implementar checkpoints para salvar o estado do modelo em diferentes estágios e utilizar visualizações, como curvas de perda e amostras de imagens geradas em intervalos regulares, para avaliar o progresso. Ajustes nos parâmetros do modelo ou na função de perda podem ser necessários com base no desempenho observado.

### **Atividade 4: Avaliação Final e Reflexão**

**Descrição:** Após o treinamento, avalie o desempenho da GAN, comparando as imagens geradas com os desenhos originais em termos de qualidade visual, diversidade e autenticidade. Prepare uma análise crítica dos resultados, discutindo as limitações observadas, os desafios enfrentados durante o treinamento e possíveis melhorias ou aplicações futuras da tecnologia.

