



# **CLASSIFICADOR DE IMAGEM**



Alunos: Paulo Costa / Leandro Rezende  
Professor: Rafael Marconi Ramos



# O QUE É UMA LLM MODAL?

Uma LLM multimodal é uma Inteligência Artificial capaz de entender e gerar não apenas texto, mas também outros tipos de dados, como:

- Imagens
- Vídeo
- Áudio
- Códigos
- Documentos
- Comandos
- Contextos mistos (texto + imagem)

Esses modelos não apenas “leem” texto – eles interpretam múltiplas modalidades de informação, por isso o nome multimodal.

# CLASSIFICADOR DE IMAGEM

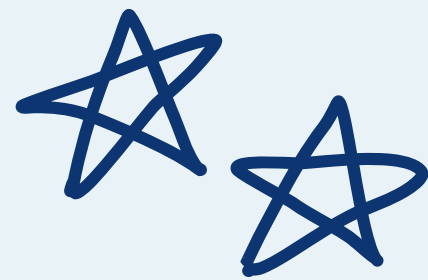
**Este projeto utiliza a tecnologia de LLMs multimodais (Large Language Models), especificamente a API Gemini, para analisar imagens e gerar classificações inteligentes.**

**Diferente de modelos tradicionais, o LLM é capaz de interpretar contexto, descrever objetos, explicar o que está vendo e fornecer respostas flexíveis, não limitadas a categorias fixas.**

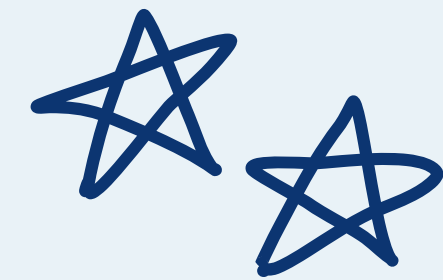
**A aplicação permite ao usuário enviar uma imagem e receber uma análise detalhada, funcionando como uma ferramenta educacional e demonstrativa de IA moderna.**

# TECNOLOGIAS UTILIZADAS

- **Interface: Streamlit**
- **LLM: Gemini 2.5 flash**
- **Linguagem: Python**
- **Bibliotecas de processamento de imagem: PIL**



# **CATEGORIAS DE CLASSIFICAÇÕES**



## **CATEGORIA 1 – Conteúdos adultos**

- Nudez
  - Seminudês
  - Atividade sexual
  - Conteúdo sugestivo
- 

## **CATEGORIA 2 – Violência**

- Sangue
  - Ferimentos
  - Armas brancas
  - Armas de fogo
  - Violência explícita
  - Violência doméstica
- 

## **CATEGORIA 5 – Discurso problemático**

- Bullying
  - Gestos ofensivos
- 

## **CATEGORIA 3 – Drogas**

- Uso de drogas ilícitas
  - Consumo de álcool
  - Consumo de cigarro/vape
  - Paraphernália de drogas
- 

## **CATEGORIA 4 – Conteúdos perigosos**

- Automutilação
  - Tentativa de suicídio
  - Comportamentos perigosos (ex.: perigo de queda)
  - Crianças em perigo
  - Conteúdo ilegal
  - Conteúdo perturbador (morte, cadáver, etc.)
-

# Vantagens e desvantagens

## VANTAGENS



- Fácil implementação
- Alta flexibilidade
- Baixo custo de manutenção
- Ideal para prototipagem rápida
- Atualizações automáticas
- Integração com várias linguagens
- Facilidade para expandir funcionalidades

## DESVANTAGENS



- Custo por tokens
- Precisão não especializada
- Limitações na interpretação visual
- Dependência da internet
- Limites de tamanho da imagem
- Variação de resposta

# DIFICULDADES

- Pouquíssimas APIs gratuitas, que suportam as requisições de imagens
- Fazer a API funcionar
- Muitas restrições de cada API, principalmente as gratuitas
- Arrumar a interface no streamlit



**THANK YOU**