



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.

Embora poderoso, o sistema não deve ser usado para decisões críticas, como identificação de alimentos venenosos, plantas tóxicas ou situações que envolvam riscos à saúde.

TECNOLOGIAS UTILIZADAS

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

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
- Riscos em decisões críticas
- 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

The slide features a light blue background with abstract blue geometric shapes and lines in the corners. In the top-left corner, there are several thin, curved lines. In the top-right corner, there is a thick, dark blue shape with a white geometric pattern. In the bottom-left corner, there is a thick, dark blue shape with a white geometric pattern. In the bottom-right corner, there are several thin, curved lines.

THANK YOU