

# Desafio Técnico – Processo Seletivo para Analista de Inteligência Artificial

## 1. Introdução e Objetivo

Este documento detalha o desafio técnico para a vaga de Analista de Inteligência Artificial. O objetivo é proporcionar uma oportunidade para você demonstrar suas habilidades práticas na construção de uma solução de ponta a ponta, desde a coleta e o tratamento de dados até a modelagem com Machine Learning e a análise de resultados.

Queremos avaliar não apenas o resultado final, mas principalmente seu processo de raciocínio, suas decisões técnicas e a clareza com que você estrutura e documenta uma solução para um problema complexo, refletindo as atividades diárias da posição.

## 2. Descrição do Desafio

Você deverá desenvolver uma solução em um notebook Python para analisar, classificar e descrever o conteúdo de uma cena simples, gravada em vídeo.

As tarefas a serem desenvolvidas são:

### a) Coleta e Pré-processamento de Dados

- **Gravação:** Grave um vídeo de **30 segundos** da superfície de uma mesa contendo **3 objetos diferentes** e facilmente identificáveis (ex: uma caneca, um mouse, um livro).
- **Narração:** Durante a gravação, descreva verbalmente em sua língua nativa os objetos que compõem a cena.
- **Separação:** Desenvolva um script para separar o arquivo de vídeo em seus canais de **imagem e áudio**.

### b) Armazenamento e Tratamento de Imagens

- **Extração de Frames:** A partir do canal de imagem, extraia um frame (imagem) a cada **0.1 segundo**.
- **Armazenamento:** Persista cada frame extraído no formato **FITS (.fits)** dentro de um **banco de dados local** (SQL ou NoSQL, à sua escolha).
- **Metadados:** Para cada imagem FITS, defina e implemente um **cabeçalho (header)** que contenha metadados que você julgar relevantes (ex: timestamp de extração, número do frame, etc.).

### c) Análise e Processamento de Imagens

- **Leitura:** Leia as imagens diretamente do banco de dados que você implementou.
- **Histograma:** Para cada imagem, calcule e plote o seu histograma.

- **Equalização:** Aplique a técnica de **equalização de histograma** em cada imagem.

#### d) Classificação com Machine Learning

- **Engenharia de Features:** Identifique, extraia e justifique as *features* (características) das imagens que permitirão a classificação dos 3 objetos.
- **Modelagem:** Escolha e treine um **algoritmo de Machine Learning** para aprender a classificar os objetos com base nas features extraídas.
- **Avaliação:** Apresente a **matriz de confusão** do modelo e analise sua performance, discutindo acertos, erros e possíveis causas.
- **Visualização:** Mostre alguns exemplos de imagens com os objetos devidamente classificados pelo seu algoritmo (ex: utilizando caixas delimitadoras ou rótulos sobre os objetos).

#### e) Análise de Áudio

- **Transcrição:** Utilize o canal de áudio extraído para converter a fala em texto (Speech-to-Text) e apresente o resultado.

#### f) Boas Práticas e Documentação

- **Estrutura:** Consolide todo o desenvolvimento em um único **notebook Python**.
- **Documentação:** Documente as principais funções e blocos de código de forma clara e concisa, utilizando **StringDocs**.

### 3. Entregáveis e Forma de Submissão

O prazo para a conclusão do desafio é de **48 horas** a partir do recebimento deste documento.

Você deverá disponibilizar sua solução completa em um **repositório público no GitHub**. Para submeter o desafio, basta **responder ao e-mail original com a URL do seu repositório**.

O repositório deve ser autoexplicativo e conter:

1. **Um README.md detalhado:** Este arquivo é o ponto de partida. Ele deve conter uma breve descrição do projeto, as tecnologias utilizadas e, fundamentalmente, as instruções claras de como clonar o repositório, configurar o ambiente virtual (`virtualenv`, `conda`, etc.) e executar a solução.
2. **O Notebook da Solução ( .ipynb ):** O arquivo contendo todo o código, as análises e as saídas visuais.
3. **Arquivo de Dependências ( requirements.txt ):** A lista de todas as bibliotecas Python necessárias para que a solução possa ser reproduzida.
4. **O Vídeo Original:** O arquivo de vídeo que você gravou. É aceitável incluir o vídeo diretamente no repositório para este desafio.

5. **Arquivo `.gitignore`** : Um arquivo `.gitignore` apropriado para projetos Python/Jupyter é recomendado.

## 4. Critérios de Avaliação

Após a submissão, agendaremos uma entrevista de revisão técnica na qual você apresentará sua solução, compartilhando a tela e nos guiando pelo seu repositório e notebook. Sua avaliação será baseada nos seguintes critérios:

- **Funcionalidade e Correção:** O código executa corretamente e atende a todos os requisitos do desafio.
- **Qualidade do Código e Boas Práticas:** Clareza, organização, eficiência e documentação do código.
- **Organização do Repositório:** A estrutura lógica do repositório, a qualidade do `README.md` e o uso de boas práticas de versionamento.
- **Raciocínio Técnico:** As justificativas para suas escolhas de ferramentas (banco de dados, bibliotecas), algoritmos e métodos de extração de features.
- **Capacidade Analítica:** A profundidade da sua interpretação dos resultados, especialmente da matriz de confusão e da performance geral do modelo.
- **Reprodutibilidade:** A facilidade para um terceiro clonar seu repositório, configurar o ambiente e reproduzir sua análise.
- **Comunicação Técnica:** Sua clareza e objetividade ao explicar o projeto e defender suas decisões durante a entrevista de revisão.

Estamos ansiosos para ver sua abordagem para este desafio. Boa sorte!