

## **Problema 1 – Aprendizado com dados estruturados**

Nos últimos anos, a Inteligência Artificial (IA) tem desempenhado um papel essencial na transformação digital de diversas áreas, como saúde, finanças, educação, e-commerce e segurança. O grande volume de dados disponíveis, aliado ao avanço das técnicas de análise de dados, tem permitido que empresas e organizações adotem modelos de IA para otimizar processos, melhorar a tomada de decisões e prever tendências com precisão.

Esses avanços, no entanto, dependem fortemente do uso de dados estruturados – ou seja, dados organizados de maneira sistemática, geralmente em tabelas, planilhas ou arquivos CSV. Quando bem tratados, esses dados podem alimentar algoritmos de aprendizado de máquina capazes de reconhecer padrões complexos e gerar soluções que vão desde a previsão de comportamentos até a automação de processos. Contudo, para que esses algoritmos sejam eficazes, é necessário um trabalho aprofundado de coleta, limpeza, análise e modelagem dos dados.

Neste contexto, o desafio proposto para as equipes será desenvolver uma solução prática de IA para um problema específico, utilizando dados estruturados disponíveis em fontes confiáveis. A escolha de um problema real e relevante, além de garantir que a base de dados seja de qualidade, será fundamental para o sucesso da solução.

### **Etapas do Desafio:**

#### **1. Escolha do Domínio e Definição do Problema**

Cada equipe deverá selecionar um domínio de aplicação para sua solução de IA. As áreas possíveis incluem:

- Saúde: Previsão de doenças, diagnósticos automáticos, otimização de tratamentos.
- E-commerce: Análise de comportamento de clientes, recomendação de produtos, otimização de estoques.
- Finanças: Previsão de risco de crédito, detecção de fraudes, otimização de investimentos.
- Segurança: Prevenção de crimes, análise de padrões de comportamento, previsão de incidentes.
- Outros domínios: Educação, transporte, marketing, entre outros.

O objetivo é que a solução de IA seja prática e viável dentro do contexto escolhido. A equipe deverá descrever o problema com clareza, definindo como a IA pode ser aplicada para oferecer uma solução inovadora e eficaz.

## 2. Coleta e Análise de Dados Estruturados

A equipe deverá buscar uma base de dados estruturada relevante para o problema que escolheram resolver. Algumas fontes recomendadas incluem:

- Kaggle: Plataforma que oferece competições de IA e conjuntos de dados públicos.
- UCI Machine Learning Repository: Repositório de bases de dados para aprendizado de máquina.
- Bases de dados governamentais ou públicas: Como dados sobre saúde, educação, segurança, crimes, entre outros.
- Outras fontes específicas: De acordo com o domínio escolhido.

Ao escolher a base de dados, a equipe deverá garantir que ela seja:

- Acessível e de fácil manuseio.
- Atualizada e relevante para o problema.
- Representativa do problema que será abordado, evitando vieses nos dados.

## 3. Processamento e Preparação dos Dados

Antes de alimentar o modelo de IA, é crucial realizar um trabalho de pré-processamento dos dados. Algumas das etapas essenciais incluem:

- Limpeza dos Dados: Remoção de valores ausentes, duplicados ou inconsistentes.
- Transformação de Dados: Conversão de variáveis categóricas em numéricas, normalização e padronização dos dados.
- Exploração e Análise Exploratória de Dados (EDA): Compreender a distribuição das variáveis, correlações entre elas, e identificar padrões e outliers que possam impactar os modelos.

## 4. Desenvolvimento do Modelo de IA

O próximo passo será a construção e treinamento do modelo de IA. A equipe deverá considerar diferentes abordagens para resolver o problema, como:

- Algoritmos de Regressão: Para problemas de previsão numérica.
- Algoritmos de Classificação: Para problemas de categorização de dados.
- Algoritmos de Clustering: Para segmentação de dados não rotulados.
- Redes Neurais e Deep Learning: Para problemas mais complexos, onde há grandes volumes de dados e padrões difíceis de capturar.

Importante: As equipes devem pensar em como otimizar a performance do modelo, seja através da escolha do algoritmo mais adequado, da seleção de features (variáveis mais relevantes para o modelo) ou pela aplicação de técnicas de regularização e otimização de hiperparâmetros. O foco é criar uma solução eficiente,

que consiga lidar com grandes volumes de dados, sem perder acurácia ou precisar de tempo excessivo para treinamento.

### 5. Avaliação e Melhoria do Modelo

A equipe deve avaliar o desempenho do modelo utilizando métricas adequadas ao problema, como:

- Acurácia, Precisão, Recall e F1-Score para problemas de classificação.
- Erro Médio Absoluto (MAE) e Raiz do Erro Quadrático Médio (RMSE) para problemas de regressão.
- AUC-ROC para avaliar a capacidade de discriminação do modelo em problemas de classificação binária.

Após a avaliação inicial, a equipe deve pensar em formas de otimizar ainda mais o modelo. Isso pode incluir técnicas como:

- Ajuste de Hiperparâmetros: Testar diferentes valores para as variáveis de configuração do modelo.
- Validação Cruzada: Para garantir que o modelo não está se sobreajustando aos dados de treinamento.
- Feature Engineering: Criar novas variáveis a partir das existentes para melhorar a precisão do modelo.

### 6. Apresentação da Solução

Após a implementação da solução, as equipes devem preparar uma apresentação detalhada que inclua:

- Descrição do problema: O que a equipe pretende resolver e qual a relevância do problema escolhido.
- Processo de coleta de dados: Detalhes sobre as bases de dados utilizadas, como foram coletadas e como foram tratadas.
- Modelo de IA: Explicação das escolhas de algoritmos, estratégias de otimização e validação.
- Resultados: Apresentação das métricas de desempenho do modelo e comparação com outras abordagens.
- Considerações sobre a solução: Limitações, possíveis melhorias e aplicação real do modelo no domínio escolhido.