

# Projeto SmartNature

## Integrantes

Angello Turano RM: 556511  
Cauã Sanches RM:558317  
Leonardo Bianchi RM:558576

## Relatório Técnico — Detecção de Risco de Queimadas com IoT e Machine Learning

### 1. Descrição do Problema

As queimadas representam um dos principais problemas ambientais no Brasil, gerando impactos climáticos, prejuízos econômicos e ameaças à vida humana e animal. A prevenção e detecção precoce desses eventos são cruciais, e a tecnologia pode ser uma aliada fundamental nesse processo.

Este projeto tem como objetivo **prever automaticamente o risco de queimadas** com base em **dados coletados por sensores IoT** instalados em regiões propensas a incêndios. Esses sensores monitoram variáveis ambientais como **temperatura, umidade do ar, nível de fumaça e velocidade do vento**.

A solução proposta visa detectar situações de risco de forma **proativa**, utilizando algoritmos de Machine Learning treinados com dados ambientais, permitindo a **emissão de alertas em tempo real** para autoridades ambientais e equipes de prevenção.

---

### 2. Metodologia Utilizada

#### 2.1. Coleta e Simulação de Dados

Como os sensores ainda não estavam disponíveis no momento do desenvolvimento, os dados foram simulados com base em padrões reais observados em estudos ambientais. Foram geradas 1.000 amostras, contendo as seguintes variáveis:

- **temperatura** (°C)

- **umidade** (%)
- **fumaca** (ppm)
- **vento** (km/h)

A variável-alvo **risco\_queimada** foi definida como binária:

- **0** = sem risco
- **1** = com risco

Essa classificação foi baseada na regra:

Se **temperatura** > 35°C e **umidade** < 30% e **fumaca** > 120, então considera-se como "risco de queimada".

---

## 2.2. Pré-processamento e Análise Exploratória

Foram aplicadas técnicas de análise exploratória de dados (EDA) para avaliar a distribuição das variáveis, detectar possíveis outliers e entender a correlação entre os atributos e a variável de risco. Também foram gerados histogramas e médias por classe.

---

## 2.3. Treinamento do Modelo

Foi utilizado o algoritmo **Random Forest Classifier**, por sua robustez, interpretabilidade e capacidade de lidar com relações não-lineares entre variáveis. As etapas do pipeline de Machine Learning foram:

- Divisão dos dados em **treino (70%)** e **teste (30%)** com estratificação da variável-alvo;
- Treinamento do modelo no conjunto de treino;
- Avaliação com métricas de classificação: **precisão, recall, F1-score e matriz de confusão**;
- Validação cruzada (5-fold) para verificar a estabilidade do modelo;
- Avaliação da **importância das variáveis** (sensores) para o modelo.

---

## 3. Resultados Obtidos

### 3.1. Métricas de Avaliação

O modelo demonstrou **boa performance**, com resultados consistentes entre treino e teste:

- **Acurácia média** na validação cruzada: aproximadamente **96%**
- **Precisão** e **recall** elevados, especialmente para a classe de risco (1)
- **F1-score** acima de 0,9, indicando equilíbrio entre precisão e recall

### 3.2. Importância dos Sensores

O modelo também indicou quais variáveis tiveram maior influência na previsão:

1. **Fumaça** → maior impacto
2. **Temperatura**
3. **Umidade**
4. **Vento** → menor impacto

Esses dados podem orientar a priorização de sensores no sistema IoT real.

---

## 4. Conclusões

- A solução proposta mostra que é **viável prever o risco de queimadas** utilizando dados ambientais monitorados em tempo real por sensores IoT.
- O modelo Random Forest mostrou-se eficaz, com **alto desempenho preditivo** e **boa capacidade de generalização**.
- A análise de importância das variáveis permite otimizar a arquitetura do sistema, focando nos sensores mais relevantes.
- O próximo passo será a **integração com os dispositivos IoT reais**, permitindo coleta de dados contínua, alimentação do modelo em tempo real e disparo automático de alertas para autoridades e brigadas.

