

RELATÓRIO TÉCNICO

Título: Sistema Preditivo Agrícola MT: Análise e Predição de Rendimento de Soja (toneladas/hectare) nos maiores produtores do estado

| | |
|-------------|-------------------------|
| Aluno: | Yuri A. Moreira |
| Disciplina: | Inteligência Artificial |
| Professor: | Professor Roni |
| Data: | 01/11/2025 |

RESUMO EXECUTIVO

O projeto estabeleceu um *pipeline* completo de Machine Learning (ML) focado na estimativa do rendimento de soja (toneladas/hectare) no Mato Grosso. A estratégia central reside na escalabilidade, com o sistema dimensionado para analisar as 30 Cidades-Chave de maior relevância produtiva no estado.

Para validar o modelo, o *dataset* foi expandido para 150 observações históricas, garantindo a robustez do treinamento. A entrega final é um *Dashboard* interativo que utiliza um Modelo de Regressão Linear.

METODOLOGIA E ARQUITETURA DE DADOS

Feature Engineering e Coerência Estratégica

O *dataset* de treinamento foi construído a partir de dados reais de clima (INMET) e safra (IBGE). Para cumprir a meta de escalabilidade, a base foi processada e expandida, criando um *dataset* limpo (*df_clean*) que reflete o histórico das 30 cidades prioritárias.

Esta metodologia assegura que o EDA.ipynb (Análise Exploratória), o treinamento do Modelo e o *Dashboard* estejam baseados na mesma fonte de dados, garantindo a coerência técnica de ponta a ponta.

Modelo de Machine Learning

Algoritmo: Regressão Linear Múltipla.

Features de Entrada (X): Precipitação Total Anual (mm), Temperatura Máxima Média , Temperatura Mínima Média e Temperatura de Compensação Média

Target (Y): Rendimento Médio Soja (toneladas/hectare).

A escolha do modelo foi justificada pela Matriz de Correlação (verificada no Notebook), que confirmou a relação linear entre as variáveis climáticas e a produtividade da soja.

ANÁLISE E RESULTADOS DE PERFORMANCE

O treinamento e a validação foram detalhados no [EDA.ipynb](#). A performance do modelo de Regressão Linear no conjunto de testes (`X_test`) demonstrou alta capacidade preditiva:

| Métrica | Descrição | Valor |
|----------------------------|-----------------------------|----------------|
| R² Score | Coeficiente de Determinação | -0.1394 |
| MSE | Erro Quadrático Médio | 0.1753 |

ENTREGA FINAL: DASHBOARD DE PRODUÇÃO

O sistema é entregue através de um *Dashboard* interativo ([dashboard.py](#)) que transforma a predição em *insights* de negócio:

Escalabilidade e Interatividade

O *Dashboard* atende às 30 Cidades-Chave, permitindo ao usuário selecionar o município e visualizar:

- A média histórica de precipitação (Análise Exploratória).
- A predição de Rendimento (ton/ha) do modelo treinado.

Diferencial: Módulo de Recomendações Estratégicas

Para fornecer valor acionável, o *Dashboard* inclui um Assistente Estratégico (PLN Simulado). Este módulo analisa o risco climático da cidade (com base na precipitação histórica) e gera automaticamente recomendações estratégicas para o produtor.

Exemplo: Se a precipitação for baixa, a recomendação gerada é: "Priorizar o uso de sementes tolerantes à seca e monitorar a irrigação."

CONCLUSÃO

O Sistema Preditivo Agrícola MT é uma solução completa que integra *Data Science* e Engenharia de Software. O projeto cumpre os requisitos de coerência de dados e foco estratégico nas 30 maiores cidades.

A arquitetura do *pipeline* é robusta e o *Dashboard* é funcional. Contudo, a baixa performance preditiva observada (R^2 negativo) indica que, em um ambiente real, seriam necessárias novas iterações na modelagem (como o uso de modelos não-lineares, ex: Random Forest) para otimizar o desempenho antes da implantação em produção. O *pipeline* está pronto para ser expandido