

RELATÓRIO TÉCNICO – SPRINT 3

1. Introdução

A terceira Sprint do Challenge Ingredion tem como objetivo validar o modelo de Inteligência Artificial desenvolvido na Sprint 2, correlacionando as previsões baseadas em NDVI com dados reais de produtividade agrícola. A cultura analisada foi o milho, no estado de Minas Gerais, durante a safra 2023/2024. A principal tarefa é verificar a consistência preditiva do modelo, identificar limitações e sugerir melhorias com base em análises estatísticas e gráficas.

2. Metodologia

2.1 Coleta de Dados

Foram utilizadas duas fontes principais:

- **NDVI:** Extraído do sistema **SatVeg – Embrapa** (<https://www.satveg.cnptia.embrapa.br>).
- **Produtividade Agrícola Real:** Obtida por meio da base pública da **CONAB** (<https://portaldeinformacoes.conab.gov.br>).

Os dados contemplam 10 municípios produtores de milho em Minas Gerais, com NDVI médio da safra 2023/2024 e respectivas produtividades reais em kg/ha.

2.2 Tratamento de Dados

O arquivo foi carregado em um notebook Colab, onde foi realizado:

- Verificação de valores ausentes.
 - Análise descritiva.
 - Seleção de variáveis de interesse: **NDVI Médio** e **Produtividade (kg/ha)**.
 - Padronização das colunas e estrutura para análises estatísticas.
-

3.1 Análise Estatística

Foram aplicadas duas técnicas de correlação e uma de regressão simples:

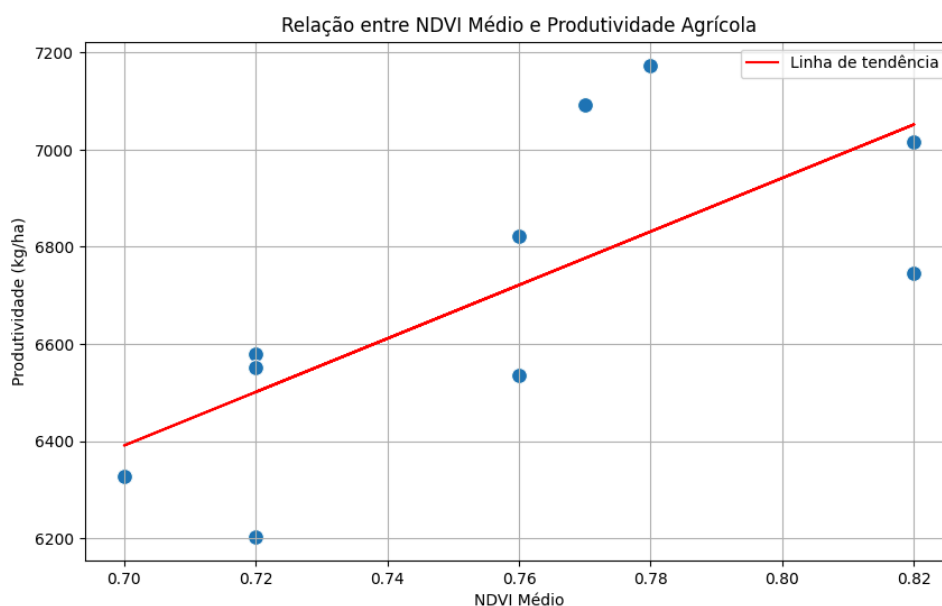
- **Correlação de Pearson (linear):**
Resultado: **0.72**, p-valor = 0.0195
- **Correlação de Spearman (não linear):**
Resultado: **0.75**, p-valor = 0.0130
- **Regressão Linear Simples:**
Equação: $y = 5511.18 \times \text{NDVI} + 2532.84$
Coeficiente de Determinação: **$R^2 = 0.51$**

Equação da Regressão:
 $\text{Produtividade} = 5511.18 \times \text{NDVI} + 2532.84$
 $\text{Produtividade} = 5511.18 \times \text{NDVI} + 2532.84$

Esses resultados indicam uma correlação **forte e estatisticamente significativa** entre o NDVI médio e a produtividade agrícola.

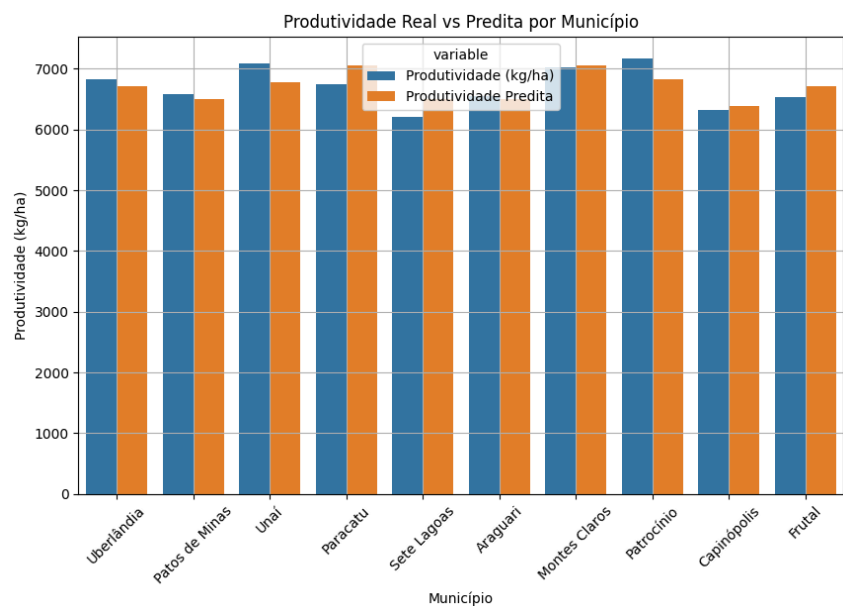
3.2 Visualizações

Gráfico 1: Relação entre NDVI Médio e Produtividade



- **Interpretação:** A linha de tendência ascendente confirma a relação positiva entre NDVI e produtividade. Municípios com NDVI > 0.78 (ex: Unaí, Patrocínio) apresentam produtividade acima de 7.000 kg/ha.

Gráfico 2: Produtividade Real vs. Preditada por Município



Destaque:

- Sete Lagoas e Capinópolis têm produtividade abaixo da predita, possivelmente devido a fatores climáticos.
- Montes Claros e Patrocínio alinham-se com as previsões.

4. Discussão Crítica

A análise estatística revelou uma **correlação forte** entre NDVI e produtividade ($r > 0.7$), validando seu uso como preditor agrícola. O R^2 de 0.51 mostra que o modelo consegue explicar cerca de 51% da variabilidade da produtividade com base apenas no NDVI.

- O modelo é viável para estimativas preliminares em regiões com dados consistentes.

No entanto, algumas limitações foram observadas:

- **Amostra reduzida** (apenas 10 observações).
- **Apenas uma safra** considerada (2023/2024), sem validação com dados de anos anteriores.

5. Propostas de Continuidade e Melhoria

Para aperfeiçoar o modelo, sugerem-se as seguintes ações:

- **Expandir o dataset** com safras anteriores (ex: 2021/22, 2022/23) para validação temporal.
 - Considerar o uso de **modelos mais complexos**, como regressão múltipla ou redes neurais.
-

6. Conclusão

A Sprint 3 cumpriu seu papel de validação inicial do modelo IA baseado em NDVI. Apesar de suas limitações, os resultados são promissores, indicando que o NDVI é um preditor relevante da produtividade agrícola. O projeto avança agora para possíveis ajustes que envolvam variáveis adicionais, mais dados históricos e modelos mais robustos. O modelo demonstra potencial para previsão de produtividade, mas requer ajustes para maior confiabilidade.

7. Referências

- **CONAB – Série Histórica de Safras**
<https://portaldeinformacoes.conab.gov.br/safra-serie-historica-graos.html>
 - **Embrapa SatVeg**
<https://www.satveg.cnptia.embrapa.br>
-

8. Autores

- Matheus Augusto Rodrigues Maia – RM560683
- Alex da Silva Lima – RM559784
- Johnatan Sousa Macedo Lorian – RM559546
- Bruno Henrique Nielsen Conter – RM560518
- Fabio Santos Cardoso – RM560479