

Análise Preditiva do Crescimento Agrícola: Abordagem Dual com Machine Learning

Problema de Pesquisa

- Prever crescimento da produção agrícola por UF (2007-2023).

Abordagem Metodológica

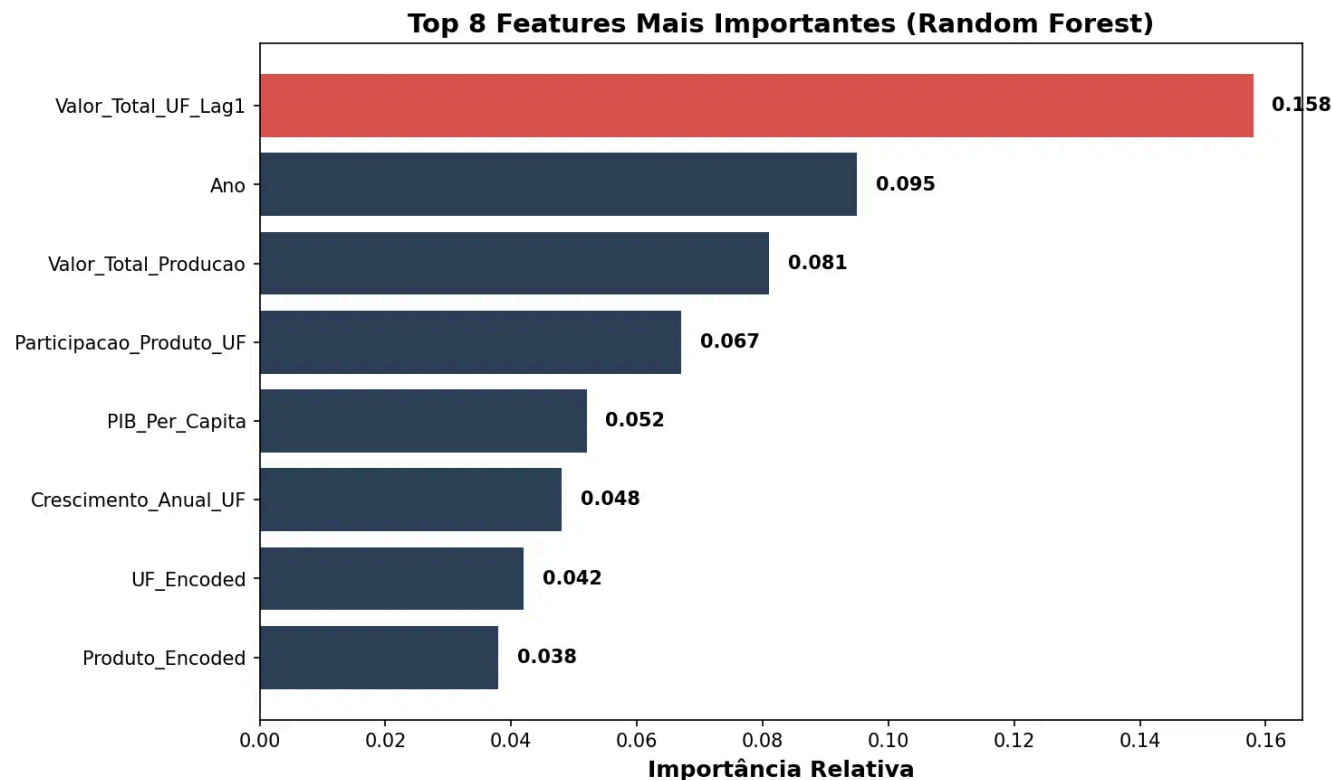
- **Classificação:** Identificar UFs com crescimento acima da média.
- **Regressão:** Prever taxa de crescimento anual.

Dados Utilizados

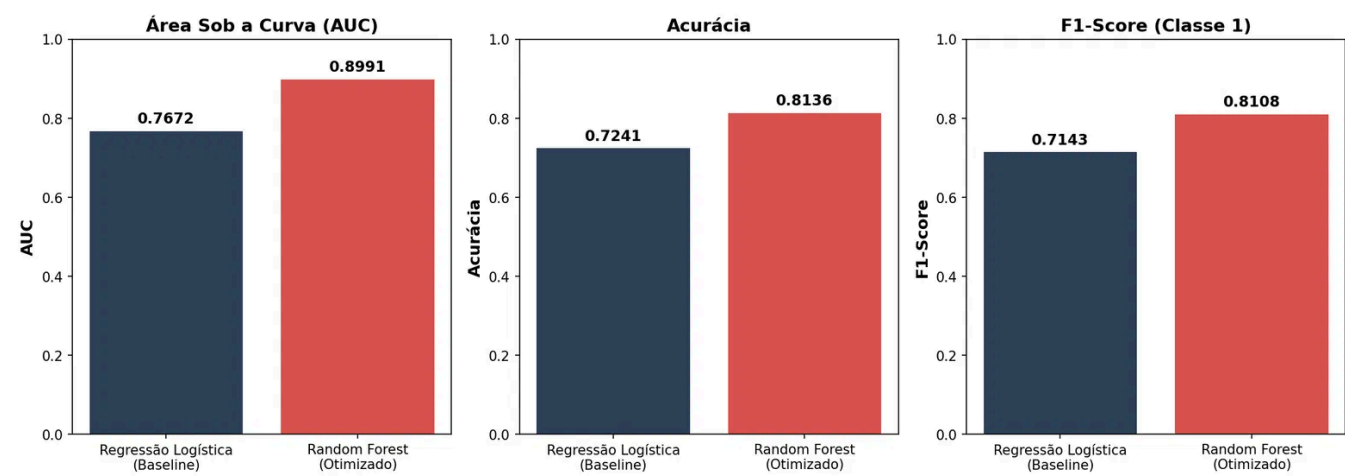
- 35 produtos, 27 UFs, 17 anos (PAM/IBGE).
- Enriquecimento com PIB per capita regional.

Modelos Implementados

- **Baseline:** Regressão Logística e Linear.
- **Avançado:** Random Forest (Otimizado).



Performance do Modelo de Classificação: Random Forest supera baseline em 17.2% (AUC)



Métrica	Baseline	Random Forest	Ganho
AUC	0.7672	0.8991	+17.2%
Acurácia	0.7241	0.8136	+12.4%
F1-Score	0.7143	0.8108	+13.5%

Interpretação

- AUC de 0.8991:** Excelente capacidade de discriminação entre classes.
- Acurácia de 81.36%:** Modelo acerta 8 em cada 10 previsões.
- Random Forest captura relações não-lineares complexas.

Modelo de Regressão alcança R^2 de 0.9578, explicando 95.78% da variância

Métrica	Baseline	Random Forest	Ganho
R^2	0.8731	0.9578	+9.7%
RMSE	0.11	0.06	-45.5%
MAE	0.07	0.02	-71.4%

Principais Insights

- R^2 de 0.9578:** Alta precisão preditiva.
- MAE de 0.02:** Erro médio de apenas 2 pontos percentuais.
- Feature #1:** Histórico de produção (Lag1) com 15.8% de importância.

