



## **Nome do projeto:**

**Modelagem de IA da FarmTech Solutions (Capítulos 13 e 14)**

## **Nome do grupo: Grupo 27**



### **Integrantes:**

- Fátima Vilela Candal – RM563003



### **Professores:**

#### **Tutor(a)**

- Leonardo Ruiz Orabona

#### **Coordenador(a)**

- André Godoi Chiovato

### **Descrição**

Analisar a base de dados (Atividade\_Cap\_14\_produtos\_agricolas.csv) com informações de condições de solo e temperatura, relacionadas com o tipo de produto agrícola e construir alguns modelos preditivos e compará-los em termos da sua performance.

## Estrutura de pastas

Os arquivos estão GITHUB no caminho:

<https://github.com/rm563003/FIAP/tree/main/FASE%203%20-%20CAP%2014/Repository>

### **FIAP / FASE 3 – CAP 14 / Repository /**

- **assets:** Imagem: VARIAVEIS.png | COMPARACAO.png
- **document:** Documentos do projeto:  
FatimaCandal\_RM563003\_fase3\_cap14.html
- **src:** Código fonte criado e arquivo csv:  
FatimaCandal\_RM563003\_fase3\_cap14.ipynb  
Atividade\_Cap\_14\_produtos\_agricolas.csv
- **README.pdf**

## Como executar o código

1 - No Jupyter Notebook fazer o Upload dos arquivos "FatimaCandal\_RM563003\_fase3\_cap14.ipynb" e "Atividade\_Cap\_14\_produtos\_agricolas.csv";

2 - Executar o arquivo "FatimaCandal\_RM563003\_fase3\_cap14.ipynb" no Jupyter Notebook .

## Comparação dos Modelos

## Comparação de Acurácias:

Os modelos exibiram as seguintes acurácias:

- Random Forest: 99%
- SVM: 98%
- KNN: 97%
- Regressão Logística: 94%
- Redes Neurais (MLP): 95%

## ## Pontos Fortes e Limitações:

- Random Forest teve a melhor performance global, sendo robusto em diferentes classes. Seu desempenho é ótimo para capturar interações complexas entre variáveis, mas pode ser mais lento e menos interpretável.
- SVM demonstrou alta precisão, especialmente em classes bem separadas. No entanto, o custo computacional pode ser elevado conforme o tamanho do conjunto de dados cresce.
- KNN teve uma acurácia relativamente alta, mas apresenta desafios em termos de eficiência para conjuntos grandes, devido à necessidade de calcular distâncias para cada nova predição.
- Regressão Logística foi o modelo com menor acurácia, possivelmente devido à linearidade dos dados. Além disso, o aviso de ConvergenceWarning sugere que ajustes nos hiperparâmetros ou pré-processamento podem ser necessários.
- Redes Neurais (MLP) mostraram desempenho robusto, ficando acima da regressão logística, mas abaixo dos modelos de árvore. Dependem de um bom ajuste de hiperparâmetros para evitar problemas como overfitting.

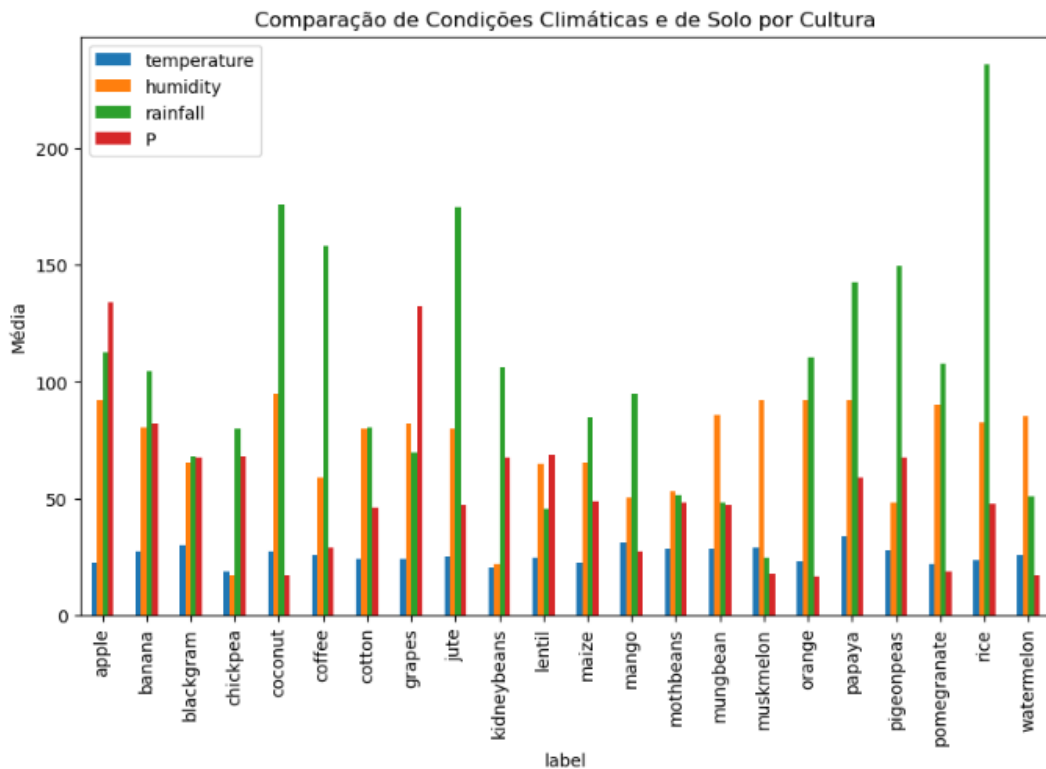
## ## Variáveis de Maior Impacto:

Embora os dados exatos das variáveis mais influentes não estejam explicitamente detalhados, modelos como Random Forest permitem análise de importância das variáveis. Em geral, fatores como clima, tipo de solo e características nutricionais tendem a ter maior peso na predição de cultivos.

## ## Perfil Ideal e Relação com Diferentes Culturas

O perfil ideal identificado tende a favorecer cultivos específicos conforme suas necessidades ambientais. Algumas observações:

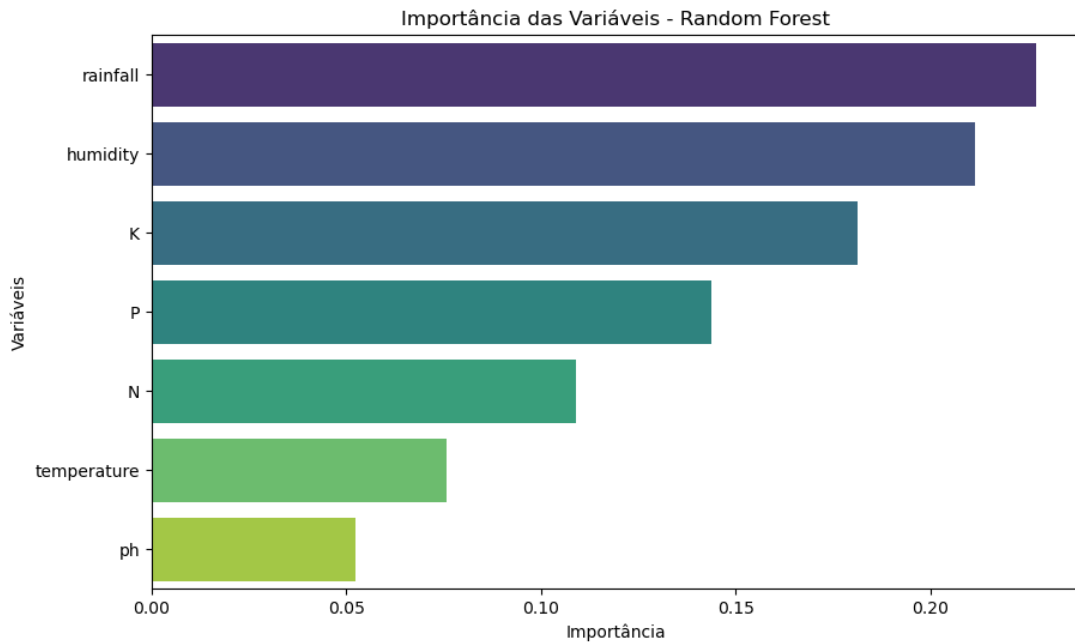
- Culturas como banana, coco e café mostram excelente classificação, sugerindo alta previsibilidade baseada nos fatores ambientais.
- Juta, lentilha e arroz apresentaram pequenas oscilações na precisão de alguns modelos, o que pode indicar maior sensibilidade a variações nos dados.
- O impacto dos fatores ambientais na previsão pode estar relacionado à regionalização da agricultura. Por exemplo, um modelo que capta padrões climáticos distintos pode ser mais eficiente na predição de cultivos em diferentes regiões.



## ## Visualizações e Comentários Adicionais:

Uma análise gráfica pode revelar mais sobre o comportamento dos modelos

- Gráfico de importância das variáveis no Random Forest.



## Conclusão

Exploramos a base de dados, realizamos análises descritivas e implementamos cinco algoritmos de Machine Learning para prever o tipo de cultura agrícola.

Sua análise permitiu compreender melhor as condições ideais de solo e clima para diferentes culturas agrícolas, além de avaliar a eficácia dos métodos preditivos.

O desempenho dos modelos variou, e o Random Forest se destacou pela sua alta precisão.

## Licença

[MODELO GIT FIAP](#) por [Fiap](#) está licenciado sobre [Attribution 4.0 International](#).