

Interpretação dos Resultados e Insights

Contexto do Problema O objetivo central deste estudo foi aplicar a metodologia CRISP-DM para automatizar a classificação de variedades de grãos de trigo (*Kama*, *Rosa* e *Canadian*), visando substituir o processo manual em cooperativas agrícolas, que é historicamente lento e propenso a falhas humanas.

Desempenho dos Algoritmos Após a fase de pré-processamento (limpeza e padronização com *StandardScaler*) e divisão dos dados (70% treino / 30% teste), foram avaliados três algoritmos distintos. Os resultados de acurácia obtidos foram:

- **K-Nearest Neighbors (KNN):** O modelo obteve o melhor desempenho, alcançando aproximadamente **100% de acurácia**.
 - *Interpretação:* O sucesso do KNN sugere que as variedades de grãos possuem fronteiras geométricas muito bem definidas. Como o algoritmo classifica baseando-se na proximidade física (distância euclidiana) dos vizinhos, ele conseguiu separar perfeitamente as classes que possuem dimensões (área, perímetro) distintas.
- **Naive Bayes (GaussianNB):** Apresentou uma acurácia sólida de **87%**.
 - *Interpretação:* Apesar da suposição "ingênua" de independência entre as variáveis, o modelo foi capaz de generalizar bem o problema. Sua ligeira desvantagem em relação ao KNN indica que as correlações entre as medidas (ex: perímetro e área são naturalmente correlacionados) são importantes para a classificação perfeita.
- **Random Forest:** Inicialmente atingiu **89%** de acurácia com parâmetros padrão.
 - *Otimização:* Após a aplicação do *GridSearchCV* para ajuste de hiperparâmetros (*n_estimators*, *max_depth*, *criterion*), o modelo apresentou uma melhoria marginal, subindo para **90%**. Isso demonstra a robustez do algoritmo, que já performa bem "out-of-the-box", mas também indica que o ganho de complexidade computacional da otimização trouxe retornos decrescentes.

Conclusão e Impacto no Negócio Os resultados evidenciam que a automatização da classificação de grãos é não apenas viável, mas altamente recomendada.

1. **Eficiência Operacional:** A implementação do modelo **KNN**, dado seu desempenho superior e simplicidade, permitiria à cooperativa processar grandes volumes de grãos instantaneamente, eliminando o gargalo da inspeção manual.
2. **Qualidade e Padronização:** O erro humano na classificação visual é eliminado, garantindo que lotes de trigo *Canadian*, por exemplo, não sejam misturados com *Kama*, o que valoriza o produto final no mercado.
3. **Recomendação Técnica:** Para implantação (Deployment), recomenda-se o uso do modelo KNN, visto que o custo computacional de inferência é baixo para este volume de dados e a precisão foi máxima.