

[← VOLTAR À LISTA](#) Fase 4 - Safras do Conhecimento Cibernetico - pragas virtuais > Cap 3 - (IR ALÉM) Implementando Al...

Cap 3 - (IR ALÉM) Implementando Algoritmos de Machine

Learning com Scikit-learn

DE 13/11/2025 A 27/11/2025

ENTREGA PENDENTE

ENTREGAR ATIVIDADE

INTRODUÇÃO

ATIVIDADE (IR ALÉM) – Da Terra ao Código: Automatizando a Classificação de Grãos com Machine Learning

Atenção Atividade Avaliativa:

- Verifique se o arquivo do upload está correto, não é possível enviar um outro arquivo após fechamento da entrega na plataforma ou correção do professor.
- Não deixe para realizar a entrega da atividade nos últimos minutos do prazo você pode ter algum problema e perder a entrega. As entregas são realizadas apenas pela plataforma.
- Não disponibilize a resposta da sua atividade em grupos de WhatsApp, Discord, Microsoft Teams, pois pode gerar plágio e zerar a atividade para todos.
- Você tem um período máximo de 15 dias após a publicação da nota para solicitar a revisão da correção.

Em cooperativas agrícolas de pequeno porte, a classificação dos grãos é realizada manualmente por especialistas, o que pode ser demorado e sujeito a erros humanos. Com o avanço do aprendizado de máquina, é possível automatizar esse processo, aumentando a eficiência e a precisão da classificação.

O objetivo é aplicar a metodologia CRISP-DM para desenvolver um modelo de aprendizado de máquina que classifique variedades de grãos de trigo com base em suas características físicas. Para isso você deverá:

- Analisar e pré-processar os dados fornecidos.
- Implementar e comparar diferentes algoritmos de classificação.
- Otimizar os modelos para melhorar o desempenho.
- Interpretar os resultados e extraer insights relevantes.

Descrição do Conjunto de Dados:

Para esta atividade, utilizaremos o "Seeds Dataset" disponível no UCI Machine Learning Repository.

Faça o download do Dataset "Seeds" em UCI Machine

Learning Repository:

<<https://archive.ics.uci.edu/dataset/236/seeds>>.

O conjunto de dados contém medições de 210 amostras de grãos de trigo pertencentes a três variedades diferentes:

1. Kama;
2. Rosa;
3. Canadian.

Atributos do conjunto de dados:

1. Área: medida da área do grão.

2. Perímetro: comprimento do contorno do grão.

3. Compacidade: calculada como $\frac{4\pi * area}{perímetro^2}$.

4. Comprimento do Núcleo: comprimento do eixo principal da elipse equivalente ao grão.

5. Largura do Núcleo: comprimento do eixo secundário da elipse.

6. Coeficiente de Assimetria: medida da assimetria do grão.

7. Comprimento do Sulco do Núcleo: comprimento do sulco central do grão.

Tarefas:

1. **Analizar e pré-processar os dados fornecidos:** nesta etapa, você deve se familiarizar com o conjunto de dados, entender os atributos e como eles se correlacionam. Siga os passos a seguir:

- Crie um arquivo notebook (.ipynb), pode ser jupyter (localmente) ou google Colab (em nuvem);
- Importe o conjunto de dados e exiba as primeiras linhas para familiarizar-se com a estrutura;
- Calcule estatísticas descritivas (média, mediana, desvio padrão) para cada característica.
- Visualize a distribuição das características utilizando histogramas e boxplots.
- Utilize gráficos de dispersão para identificar possíveis relações entre as características.
- Identifique e trate valores ausentes;
- Avalie a necessidade de escalar as características e aplique normalização ou padronização se necessário.

2. **Implementar e comparar diferentes algoritmos de classificação:** nesta etapa, você deverá construir modelos de classificação utilizando diferentes algoritmos e comparar seus desempenhos. Siga os seguintes passos:

- Separe os dados em conjuntos de treinamento e teste (por exemplo, 70% para treinamento e 30% para teste).
- Escolha pelo menos três algoritmos de classificação diferentes. Por exemplo:

- K-Nearest Neighbors (KNN);
- Support Vector Machine (SVM);
- Random Forest;
- Naive Bayes;
- Logistic Regression.

• Treine cada modelo utilizando o conjunto de treinamento.

• Avalie cada modelo no conjunto de teste. Use métricas de desempenho, como acurácia, precisão, recall, F1-score e matrizes de confusão.

• Compare o desempenho dos diferentes algoritmos.

3. **Otimizar os modelos para melhorar o desempenho (se necessário):** após a avaliação inicial, você deve avaliar se é necessário otimizar os modelos para aprimorar o desempenho, caso seja necessário:

- Utilize Grid Search ou Randomized Search para encontrar os melhores hiperparâmetros para cada modelo.
- Treine novamente cada modelo utilizando os melhores hiperparâmetros encontrados.
- Avalie novamente cada modelo otimizado no conjunto de teste. Use métricas de desempenho, como acurácia, precisão, recall, F1-score e matrizes de confusão.
- Verifique se houve melhorias significativas no desempenho.

4. **Interpretar os resultados e extraír insights relevantes:** por fim, você deverá analisar profundamente os resultados e extraír conclusões significativas. Interprete o desempenho de cada modelo e relate os resultados com o contexto do nosso problema de classificação de grãos.

Organize sua atividade no GitHub como FASE 04/CTWP/Cap3, utilizando o modelo solicitado pelo tutor, mande o link no portal do aluno e não mexa no repositório após a data de vencimento.

GRUPO

Mínimo: 1 | Máximo: 5

Participante (s)

Procure por nomes ou RM dos participantes

<input checked="" type="checkbox"/> Viviane de Castro vivi.topproducer@gmail.com	X
<input checked="" type="checkbox"/> GUILHERME FERREIRA SANTOS gifikasi.channel@gmail.com	X
<input checked="" type="checkbox"/> André Pessoa Gaidzakian andrepgaidzak@gmail.com	X
<input checked="" type="checkbox"/> Erick Prados Pereira erick.prados.pereira@gmail.com	X

Envio

ENTREGA PENDENTE

Avaliação

AVALIAÇÃO PENDENTE

PRAZO DE ENTREGA

Quinta-feira, 27 de Novembro de 2025, às 23h59

TEMPO RESTANTE

4 dias e 8 horas restantes

ÚLTIMA MODIFICAÇÃO

-

GRUPO

Grupo 4

INTEGRANTES DO GRUPO 4

Viviane de Castro
vivi.topproducer@gmail.com

GUILHERME FERREIRA SANTOS
gifikasi.channel@gmail.com

André Pessoa Gaidzakian
andrepgaidzak@gmail.com

Erick Prados Pereira
erick.prados.pereira@gmail.com