



Matplotlib

Semana 6 / Atividade 3

Nesta atividade, você vai aprender Matplotlib, a base de visualização em Python. A ideia é transformar dados e resultados (CSV de previsões e métricas de treino) em gráficos que ajudem a interpretar desempenho, erros e confiança do modelo.

Por que estudar Matplotlib se o foco é visão computacional?

Você vai precisar visualizar:

- Curvas de loss/accuracy (para diagnosticar overfitting/underfitting);
- Distribuições de probabilidades;
- Erros do modelo;
- Matrizes de confusão e comparações entre modelos.

Objetivos

Ao final deste estudo dirigido, você deverá:

- Criar gráficos simples (barras, histograma, scatter).
- Controlar título, eixos, legendas e layout.
- Visualizar distribuições de probabilidades do modelo.
- Comparar contagens reais vs previsões.
- Plotar curvas de treino/val a partir de um `metrics.csv`.

Atividades

Acompanhar as aulas e codar junto

Assista à playlist, faça exemplos no seu editor e salve no diretório src/07-matplotlib:

https://www.youtube.com/playlist?list=PL5jigOsyxDtBEgO2zUUyf_2LNt3Zenrl6

Exercícios Práticos

Dentro do diretório src/07-matplotlib/exercicios, implemente:

ex01.py – Visualização do CSV de classificação

Arquivo: classification_results_trial_0001.csv disponível no diretório de arquivos compartilhados do onboarding.

Faça os gráficos:

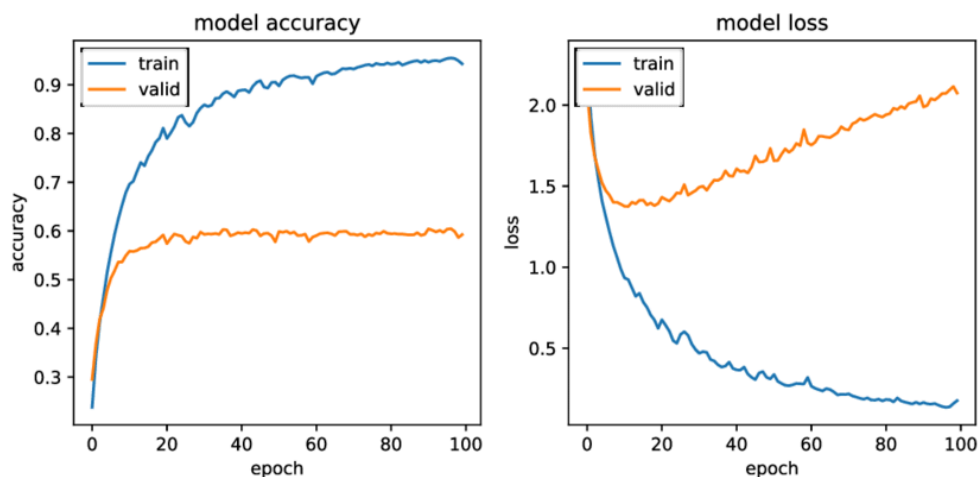
- Gráfico de barras: contagem por real_class.
- Gráfico de barras: contagem por predicted_class.
- Histograma de prob_benign.
- Histograma de prob_malign.
- Scatter plot: X=prob_benign, Y=prob_malign, com cor diferente para acerto vs erro.
- Qual erro é mais comum (FP ou FN)? Crie um gráfico que ajude a visualizar isso.
- E em contexto médico: qual é mais preocupante e por quê?

ex02.py – Curvas de treino e validação (accuracy e loss)

Arquivo: metrics.csv disponível no diretório de arquivos compartilhados do onboarding.

Contexto: Ao treinar modelos de Machine Learning ou Deep Learning, é fundamental monitorar como o modelo aprende. As métricas de loss (perda) e accuracy (acurácia) são registradas tanto para o conjunto de treinamento quanto para o conjunto de validação em cada época (um ciclo completo de treinamento sobre os dados). Analisar essas tendências nos ajuda a entender se o modelo está aprendendo bem, se está se ajustando demais aos dados de treinamento (overfitting) ou se não está aprendendo o suficiente (underfitting).

Referência de Gráfico:



Gere um gráfico semelhante ao de referência com as curvas de train_loss e val_loss e train_acc e val_acc. Use o eixo x como épocas. Use o arquivo de dados metrics.csv.

Entrega da Atividade

Gerar um arquivo PDF que contenha:

- Seu nome.
- O link do seu repositório no Github.
- O código de todas as atividades desse roteiro de estudos (código das videoaulas, reflexões e exercícios)

Enviar o arquivo PDF na tarefa S6.A3. Exercícios Matplotlib.