

Relatório final de experimentação de modelos de machine learning (8 pontos)

Cada grupo deverá entregar um relatório escrito apresentando o ciclo completo de experimentação de modelos de aprendizagem de máquina para o seu dataset. O trabalho integra a etapa exploratória realizada no Orange com a implementação dos modelos em Python, utilizando os algoritmos tradicionais trabalhados em aula (rede neural, SVM, kNN, árvore de decisão, floresta randômica, regressão logística e regressão linear, quando aplicável). O objetivo é documentar de forma estruturada o processo de escolha do modelo considerado mais adequado, justificando essa decisão com base em métricas e decisões de projeto.

A primeira etapa do relatório consiste em descrever o problema e o conjunto de dados utilizado. O grupo deve indicar se se trata de uma tarefa de classificação ou regressão, qual é a variável alvo e quais são, em linhas gerais, as principais variáveis preditoras. Também deve ser apresentada, em texto corrido, a motivação prática do problema: em que contexto esse modelo seria útil e que tipo de decisão ele apoiaria.

Na sequência, o grupo deve apresentar uma síntese da etapa de análise realizada no Orange. Essa síntese deve retomar, de forma compacta, quais modelos foram testados no ambiente visual, como foi configurada a validação e quais métricas foram usadas para comparação. É importante registrar qual modelo foi apontado como melhor nessa fase e por quais razões, tomando como referência as métricas observadas e as características do problema. Não é necessário reproduzir todo o relatório da atividade de 2 pontos, apenas o resumo que conecta essa etapa com a implementação em código.

A etapa seguinte do trabalho descreve a implementação dos modelos em Python. O grupo deve explicar que ambiente utilizou (Python, bibliotecas como scikit-learn, pandas e outras necessárias) e qual foi o pipeline básico de preparação dos dados. Nesse ponto, espera-se uma breve explicação sobre tratamento de valores ausentes, normalização ou padronização, codificação de variáveis categóricas e divisão entre dados de treino e teste ou configuração da validação cruzada.

Ainda na parte de implementação, o grupo deve informar quais modelos foram efetivamente testados em Python. A lista de candidatos pode incluir, conforme o tipo de problema, alguns ou todos os seguintes algoritmos: regressão logística, SVM, kNN, árvore de decisão, floresta randômica, rede neural e regressão linear (apenas para tarefas de regressão).

Para cada modelo, é suficiente descrever em texto os principais hiperparâmetros utilizados e o esquema de validação adotado, sem necessidade de copiar código-fonte no relatório.

Os resultados obtidos em Python devem ser organizados em uma tabela comparativa de métricas. Para problemas de classificação, recomenda-se incluir pelo menos acurácia e F1, podendo acrescentar outras métricas trabalhadas em aula, como AUC, precisão e revocação. Para problemas de regressão, recomenda-se apresentar RMSE, MAE e R². Após a tabela, o grupo deve comentar os resultados, destacando quais modelos se destacaram em cada métrica e se há trade-offs relevantes entre desempenho e simplicidade do modelo.

Uma parte central da avaliação é a comparação entre o que foi observado no Orange e o que foi obtido na implementação em Python. O grupo deve explicar se o modelo que parecia ser o melhor no Orange se manteve como melhor quando todos os modelos foram testados em código. Em caso de convergência, basta registrar essa consistência e comentar brevemente por que isso fortalece a escolha. Em caso de divergência, o grupo deve discutir possíveis motivos, como diferenças de pré-processamento, configuração de validação, hiperparâmetros padrão ou variações decorrentes de aleatoriedade.

Com base em todas as análises anteriores, o grupo deve então formular a justificativa final de escolha do modelo. Essa justificativa deve articular métricas, características do problema e decisões de projeto. É importante tratar também aspectos como interpretabilidade, complexidade do modelo e adequação prática ao cenário descrito no início do relatório. Modelos descartados devem ser mencionados de forma sintética, explicando por que foram considerados menos adequados, mesmo que tenham tido desempenho próximo em alguma métrica.

O relatório deve incluir uma breve seção sobre limitações e recomendações. Nessa parte, o grupo deve reconhecer eventuais limitações dos experimentos, como ausência de busca sistemática de hiperparâmetros, tamanho reduzido do dataset, desequilíbrio de classes ou simplificações adotadas por restrição de tempo. Também deve apontar sugestões de continuação do trabalho, como ajustes de hiperparâmetros, uso de técnicas de balanceamento, ampliação de dados ou teste de outros algoritmos.

Formato e entrega:

- Forma: relatório em PDF.
- Estrutura mínima recomendada:
 1. Identificação do grupo e do dataset.
 2. Descrição do problema e do conjunto de dados.
 3. Síntese da etapa no Orange.
 4. Implementação dos modelos em Python.
 5. Resultados e comparação entre modelos em Python.
 6. Comparativo Orange x Python.
 7. Justificativa final do modelo escolhido.
 8. Limitações e recomendações.
 - Extensão sugerida: entre 5 e 10 páginas, incluindo tabelas e figuras.
 - Nome do arquivo: ML_RelatorioFinal_GrupoX.pdf, substituindo X pelo identificador do grupo.
 - Canal e prazo de entrega: definidos no ambiente virtual da disciplina.

Critérios gerais de avaliação (total de 8 pontos)

- Clareza na descrição do problema e do dataset.
- Coerência na seleção e documentação dos modelos em Python.
- Qualidade da apresentação e interpretação das métricas.
- Profundidade da comparação Orange x Python, com explicações plausíveis para convergências ou divergências.
- Consistência e fundamentação da escolha final do modelo.
- Organização, objetividade e qualidade geral da redação técnica.