

Trabalho Prático 1

Trabalho em dupla.

Data da entrega final: 21/05/2025- Ver entregas parciais na tabela 1 - **Valor:** 15 pontos

Cada dupla deve escolher um problema no Kaggle. Elaborar uma solução e apresentar para turma no dia da entrega. A solução da dupla deve ser entregue no Moodle em um Python Notebook. Especificamente, a dupla deve:

- Pesquisar um problema **interessante** no Kaggle que envolva dados tabulares/estruturados;
- Realizar uma análise exploratória dos dados da competição. Isto é, visualizar os dados, tentar identificar dados faltantes, instâncias ruidosas e tendências nos dados;
- Entender e implementar a métrica de avaliação da competição;
- Implementar uma solução simples para o problema da competição para gerar um baseline (isto é, o desempenho mínimo que o sua abordagem deve alcançar);
- Estudar soluções para problemas semelhantes ao da competição escolhida. A dupla pode pesquisar, por exemplo, em artigos acadêmicos, artigos técnicos da Internet e o no próprio Kaggle;
- Implementar diferentes abordagens para a competição e reportar o desempenho de cada abordagem implementada;
- Realizar uma avaliação de parâmetro (*hyperparameter tuning*) das abordagens escolhidas para o problema. Para realizar a avaliação dos parâmetros você pode utilizar [GridSearchCV](#) ou [Optuna](#).
- Além de outras abordagens, a dupla deve:
 - Propor pelo menos uma abordagem que envolva *meta-learning*. Ou seja, a dupla deve propor uma forma diferente de combinar vários modelos. Por exemplo, utilize a abordagem do Random Forest, porém com vários classificadores ou invés de utilizar apenas Árvores de Decisão.
- Para cada abordagem apresentada, a dupla deve apresentar:
 - O desempenho no conjunto de treino, o desempenho deve ser medido com uma validação cruzada com 10 folds (*10-fold cross-validation*).
 - O desempenho no conjunto de teste reportado pelo Kaggle.

Atividade	Data de Entrega
Definição do conjunto de dados. Enviar nomes da dupla e link para o conjunto de dados no Kaggle no fórum no SIGAA	07/05/2025
Apresentação parcial dos resultados	14/05/2025
Apresentação final	21/05/2025

Tabela 1. Entregas