**Data da entrega:**

1. Escolha uma base de dados de classificação. O [Scikit-learn](https://scikit-learn.org/stable/datasets/toy_dataset.html) e o [UCI Repository](https://archive.ics.uci.edu/ml/index.php) possuem algumas bases. Caso algum dado seja categórico, transforme para númerico.
2. Realize uma Análise Exploratória dos Dados no conjunto de dados selecionado.
3. Divida os dados em [treino/teste](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html) utilizando a técnica hold-out. Escolha uma porcentagem para treino/teste (mais utilizadas: 90/10, 80/20, 75/25, 70/30, 50/50).
4. Realize a classificação utilizando o KNN, utilizando a biblioteca [Scikit-Learn](https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html).
5. Calcule algumas métricas. Obrigatório: acurácia. Exemplo de outras métricas: recall, precision, F1-Score. Utilize o Scikit-Learn (classification\_report, accuracy\_score, f1\_score, etc)
6. Plot a [matriz de confusão](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.ConfusionMatrixDisplay.html#sklearn.metrics.ConfusionMatrixDisplay). Outra biblioteca interessante para o plot é a seaborn ([exemplo de uso](https://medium.com/@dtuk81/confusion-matrix-visualization-fc31e3f30fea)).
7. Realize vários experimentos, modificando a quantidade de vizinhos (parametro n\_neighbors), e crie um plot (k vs acurácia).
8. Discuta os resultados encontrados.
9. Utilize a técnica k-fold e discuta os resultados, comparando com a técnica hold-out.
10. Realize uma combinação de técnicas ( exemplos: hold out + kfold, dois k-fold, entre outros) e compare os resultados com as demais questões.