

# Atividade 02

INF 2137 - Engenharia de software para ciência de dados

**Pedro Henrique Cardoso Paulo**

pedrorjpaulo.phcp@gmail.com

Professor: Marcos Kalinowski



Departamento de Engenharia Mecânica  
PUC-RJ Pontifícia Universidade Católica do Rio de Janeiro  
julho, 2023

# Atividade 02

## INF 2137 - Engenharia de software para ciência de dados

Pedro Henrique Cardoso Paulo

julho, 2023

### 1 Introdução

Esta tarefa tem por objetivo realizar um pequeno projeto de software inteligente usando as boas práticas aprendidas no curso.

### 2 Enunciado

#### 2.1 Tarefa

Escolha uma base de dados para um problema de classificação ou regressão. Sugere-se usar uma das bases de dados disponibilizada no UCI Machine Learning Repository ou no Kaggle. Você deverá trabalhar desde o problema até os resultados e necessariamente comparar e otimizar os hiperparâmetros para mais de um modelo.

Lembre-se das etapas para implementação de projetos de ciência de dados, apresentadas em sala de aula, e trabalhe com esta base de dados no ambiente Google Colab, utilizando Python e a biblioteca Scikit-Learn. Produza um notebook que servirá como relatório, descrevendo textualmente cada uma das etapas do seu código. Para a definição do problema, descreva uma user story e seu detalhamento (simplificado – não é necessário utilizar o template de especificação da atividade anterior, já que é um toy problem e não um projeto industrial), na primeira célula do notebook. Para a implementação, sugere-se utilizar as boas práticas de programação orientada a objetos trabalhadas em sala.

Na última célula, informe como você implantaria a solução de forma a disponibilizar o resultado do seu modelo para o usuário final.

### 3 Solução

#### 3.1 Problema e Dataset

O problema escolhido foi previsão de vazão de óleo em um poço de petróleo usando dados do Volve Dataset.

#### 3.2 User Story

**Como** engenheiro de reservatório **quero** prever a produção de óleo de um poço baseado em seu histórico **para** tomar decisões de priorização e cumprir obrigações legais.

#### 3.3 Link

- [Pasta geral com o conteúdo e dataset](#)
- [Notebook com entrega final](#)