

Projeto Integrador

Aprendizagem Estatística de Máquina I

O objetivo desse projeto é desenvolver um modelo preditivo para indicar se uma empresa deixará de operar em um período de até dois anos.

O pipeline de pré-processamento deverá ser feito em Python e a modelagem em R. Para a modelagem, considere as seguintes etapas:

1. Defina as métricas utilizadas para avaliar os desempenhos dos modelos.
Comente a importância dessas métricas para o problema em questão.
2. Considere, pelo menos, quatro classes de modelos para avaliação. Quando necessário, faça o ajuste dos hiperparâmetros.
3. Compare os modelos em relação ao desempenho preditivo, considerando as métricas definidas no item 1.
4. A partir do modelo selecionado, considere pelo menos uma ferramenta (VIP, PDP, lime ou shapley values) para identificar importância e/ou interpretação das preditoras consideradas.

A entrega deverá ser feita com o banco de dados utilizado na modelagem, notebook com o processamento em Python e relatório em R markdown (junto com arquivo fonte). É importante ressaltar que o relatório e o código devem ser compreensíveis para alguém que não participou de todo o processo.

Lembre-se que o processo de desenvolvimento de um modelo preditivo é um ciclo e poderá ser necessário voltar para o pré-processamento diversas vezes. Aproveite o projeto para aprofundar seu conhecimento e praticar a utilização de diferentes ferramentas!