



Especificação do Projeto 2

Luciano Barbosa



Passos

1. Escolher uma das colunas dos dados utilizados no projeto 1 para predição (classificação ou regressão)
2. Separar os dados em treinamento, validação e teste
3. Selecionar 4 algoritmos de acordo com a tarefa escolhida no passo 1
4. Adicionar MLFlow no treinamento dos modelos para rastreamento
5. Executar uma ferramenta de seleção de hiper-parâmetros sobre o conjunto de validação
 1. Selecionar poucos hiper-parâmetros por algoritmo (max. 3)
 2. Selecionar o modelo (algoritmo+valores de hiper-parâmetros) com melhor resultado na métrica de avaliação
 3. Executar o melhor modelo de cada algoritmo no conjunto de teste e selecionar o melhor modelo geral
6. Realizar diagnóstico do melhor modelo geral da etapa 5 e melhorá-lo a partir do diagnóstico
 - **Bônus:** Utilizar auto-sklearn, clustering para entendimento dos dados