

Laboratório Aprendizagem de Máquina

Lab: Regressão

Considere a base em anexo com os dados de uma usina hidroelétrica, que tem as seguintes características.

- f3) Vazão Turbinada
- f4) Vazão Vertida
- f5) Afluência
- f6) Afluência Soma Montante
- f7) Volume Prévio
- f8) Volume Atual
- f9) Volume Prévio Soma Montante
- f10) Volume Atual Soma Montante
- f11) Volume Prévio Soma Jusante
- f12) Volume Atual Soma Jusante

As características [f5..f12] devem ser utilizadas para fazer a predição das variáveis f3 e f4. Ou seja, você deve treinar dois regressores, um para cada tarefa. Note que nem todas as características precisam ser utilizadas pelo regressor.

Dentre os regressores abaixo, reporte aquele que minimiza o erro quadrado médio (MSE).

- Regressão linear
- SVR
- Redes Neurais (MLP) (+ de uma camada escondida)
- kNN
- Árvore de Decisão
- Random Forest
- Gradient Boosting

Divida a base em dois subconjuntos (50-50). Utilize o parâmetro `random_state` na função `train_test_split` para poder comparar os resultados. Para encontrar parâmetros dos regressores, você pode utilizar qualquer método de validação. 50% da base deve ser utilizado exclusivamente para testar o modelo.

Faça uma análise de resíduos e verifique se 95% dos resíduos estão no intervalo de um desvio padrão.