

# Exercício 5

**Data de entrega: 14/11, as 7:00 (da manhã).**

Os dados [train](#) representam um problema de regressão. A 1a coluna é o valor a ser aproximado.

Use pelos menos 2 técnicas para fazer a regressão ( SVM regressão, gbm, rf, redes neurais, cubist, e gaussian regression, e outras mesmo que não tenhamos visto em aula). A metrica sera MAE - erro absoluto médio (nao erro quadrado!). Reporte os resultados de pelo menos essas duas tecnicas - quais hyperparametros foram tentados e qual a acuracia da tecnica para algum valor de crosvalidação externa

Considere fazer preprocessamento dos dados, normalizacao, PCA, etc. Tambem considere reduzir o numero de classes de algumas variavies (agrupando varias classes numa so). Considere eliminar variaveis numéricas que tenham pouca variancia.

O relatório com pelo menos 2 das tecnicas tentadas (mais o presprocessamento, se foi o caso) valerá 60% da nota.

Rode o seu melhor regressor [nestes dados, e submeta também o resultado no formato do valor previsto, um por linha na mesma ordem dos dados. Note que a primeira coluna dos dados de teste corresponde a 2a coluna dos dados de treino.](#)

[Eu avaliarei o MAE do seu regressor nos resultados corretos. 40% restante da nota sera competitiva: as submissoes no topo 10% com menos MAE receberão o 10 nessa parte e as submissões nos ultimos 10% \(maiores MAE\) receberão 0.](#)

---

[Last modified: Sun Oct 23 12:59:18 BRST 2016](#)