

Análise de Regressão

Prova 1 - Parte 2

Valor 30% da nota final da prova

Data de entrega: **06 de agosto**.

Descarregue os dados *Estudantes.csv* usando o link <https://www.dropbox.com/scl/fi/c84nw3do7fk98vc36qb3e/Estudantes.csv?rlkey=z89gphri2gfvfuubjxr3vokgw&dl=0>. Estes dados são os resultados de um estudo feito a estudantes de duas escolas de ensino médio. As variáveis dos dados são:

1. Tipo: Tipo de escola. 1 Pública e 2 Privada.
2. Idade: Idade do estudante em anos.
3. TempoE: Tempo de estudo semanal, em horas, fora de aula.
4. Hsono: Número médio de horas que dorme por noite.
5. Alcool: Número médio de vezes que tem contato com álcool, por mês.
6. Depor: Número médio de horas dedicadas a alguma atividade esportiva, por semana.
7. Nota: Nota média obtida no ano de todas as disciplinas.

Fonte dos dados: Center for Machine Learning and Intelligent Systems, <http://cml.ics.uci.edu/>.

Exercícios

1. Que tipo de variáveis temos nos dados?
2. Como utilizaria a variável tipo de escola dentro da análise destes dados?
3. Apresente uma equação para um modelo de regressão linear cuja variável resposta seja a Nota média (tenha em conta que o modelo de regressão é para variáveis numéricas).
4. Apresente testes de hipóteses sobre o modelo e suas conclusões.
5. Verifique todos os supostos do modelo. O que pode concluir?
6. Apresente o código usado.