

## Nota 2 - Regressão

O objetivo desse trabalho é treinar e avaliar um modelo de regressão linear para prever o preço de produtos a partir da sua quantidade de reviews e a nota média obtida nesses reviews.

### Tarefas

- Ler arquivo com os dados de treinamento
- Treinar um modelo de regressão linear
- Avaliar o modelo usando validação cruzada
- Apresentar gráfico com pontos representando os exemplos e a reta resultado da regressão linear (bônus: 1 ponto extra)

### Arquivo

O arquivo com os dados de treinamento é o `items.csv`. Ele é um arquivo csv (comma separated values) que contém informações sobre reviews de produtos na Amazon. O arquivo possui 792 exemplos com as seguintes colunas:

- `asin`: id do produto
- `brand`: marca do produto
- `title`: nome do produto
- `url`: URL do produto
- `image`: URL da imagem do produto
- `rating`: nota média dos reviews
- `reviewUrl`: URL da página de review
- `totalReviews`: quantidade de reviews
- `prices`: preço do produto

A primeira linha do arquivo especifica o nome das colunas e as linhas seguintes os dados sobre reviews dos produtos. Cada coluna é separada por vírgula e cada linha seguinte possui os dados de um produto.

Para este trabalho os valores das colunas `rating` e `totalReviews` serão usados para prever os valores da coluna `prices`.

No arquivo, alguns exemplos não possuem preço, por isso, eles devem ser descartados.

### Resposta

O modelo deve ser avaliado utilizando validação cruzada. Assim, o programa deve exibir o erro médio quadrático no conjunto de treinamento e no conjunto de testes.

## FAQ

1. Em qual linguagem eu posso implementar o algoritmo? Java ou Python.
2. Posso usar alguma biblioteca que implementa a regressão? Sim. Recomendo a scikit-learn.
3. É necessário acertar todas as previsões? Não. Muitas vezes os modelos não conseguem acertar 100% e isso pode acontecer nesse problema também.