



EXERCÍCIO 4

REGRESSÃO LINEAR

Objetivo

Neste exercício, iremos implementar o método de Regressão Linear: um tradicional método de regressão, com conceitos que servirão de base para diversos outros métodos vistos na disciplina.

A Regressão Linear é um método de aprendizado supervisionado baseado em otimização, que tem como objetivo prever números contínuos e busca encontrar um vetor gradiente que minimize a diferença entre os valores previstos pelo método e os valores reais. O método pode ser considerado a base para diversos outros métodos que serão vistos ao longo da disciplina, como Regressão Logística e Redes Neurais Artificiais. Seus conceitos de cálculo de uma função-custo e otimização do gradiente são fundamentais para boa parte da teoria por trás de Aprendizado de Máquina.

Ao término desse exercício, espera-se que você consiga implementar o método de Regressão Linear, através do cálculo da função custo e do aprendizado por meio da técnica de gradiente descendente. Um problema de previsão orçamentária será utilizado para ilustrar o exercício e exemplificar a utilização do método.

O exercício

Ao longo do exercício, você deverá completar 3 funções:

- **plotarDados**: responsável por gerar um gráfico que permita a visualização dos dados utilizados no exercício;
- **computarCusto**: responsável por calcular o custo de usar um determinado valor de θ para o conjunto de dados do treinamento;
- **gradienteDescendente**: responsável por executar o gradiente descendente, um clássico otimizador que irá aprender o conteúdo de θ , permitindo assim a previsão;

Preencha o código apenas nos espaços delimitados por comentários, normalmente iniciados por um comentário “COMPLETE O CÓDIGO AQUI” e instruções para a implementação.

As implementações devem ser o mais genéricas possíveis, funcionando para qualquer conjunto de dados. Durante a avaliação, as funções serão testadas em bases com quantidade de amostras e atributos diferentes das fornecidas com o exercício, tendo em comum apenas o nome da coluna que contém a classe das amostras. Não adicione comandos do tipo `print` ou `display` dentro das funções que serão completadas, apenas o código da função..

Aproveite para interagir com o *notebook*: teste valores diferentes para as amostras que serão previstas e entenda como o método funciona.

Os casos de teste

Este exercício possui **5 casos de teste** que buscam avaliar cada uma das funções implementadas. A distribuição de tarefas avaliadas por cada caso é feita da seguinte forma:

- **Casos de teste 1 e 2:** corrige a função `computarCusto`;
- **Caso de teste 3 a 5:** corrige a função `gradienteDescente`.