

Proposta - Trabalho Prático 4

Análise das características de um carro para predizer o seu preço de varejo estimado.

base de dados escolhida foi encontrada no Kaggle, denominada como Car Features and MSRP (<https://www.kaggle.com/CooperUnion/cardataset>). A base de dados está no formato CSV e é composta por 16 atributos e 11915 registros diferentes, fornecidos pelos websites Edmunds e Twitter, que descrevem um carro. Dentre os atributos, temos a montadora, o modelo do carro, o ano de fabricação, o tipo do motor de gasolina, o tipo de transmissão, o número de portas, o tamanho do carro e entre outros. O dataset foi escolhido por possuir uma documentação clara e com um bom número de registros e atributos de fácil entendimento, facilitando o treinamento da rede neural para a aplicação das técnicas de regressão.

Objetivo

O objetivo deste trabalho é treinar um modelo utilizando os algoritmos de regressão, como a regressão linear, gradient boosted tree etc. Para isso, a base de dados será separada em um conjunto de treino e outro de teste e será treinado para que consiga compreender as 16 características e o rótulo msrp, que é o objetivo da predição.