

Exercício de apoio - Semana 4: Classificação supervisionada

Use árvores de decisão para classificar o conjunto de dados sobre aceitabilidade de carros de acordo com as suas características. Os detalhes do conjunto de dados estão descritos abaixo:

1. Separe os dados em treinamento e teste, e meça a acurácia do classificador. Para o treino, separe 20% dados.
2. Gere a imagem da árvore de decisão obtida no treinamento. Você pode salvar a imagem clicando com o botão direito do mouse e abrindo em um leitor de imagens para ver mais detalhes da árvore gerada.

URL do conjunto de dados

<http://archive.ics.uci.edu/ml/datasets/Car+Evaluation>

Fonte: UCI Machine Learning Repository, Centro para Aprendizado de Máquina e Sistemas Inteligentes, Universidade da Califórnia, Irvine.

*O formato do arquivo (extensão .data) é texto separado por vírgulas, então pode ser salvo como CSV.

Descrição dos atributos e seus valores:

São seis atributos de características e mais um contendo a classe dos objetos.

- **buying** (v-high, high, med, low): preço de compra (muito alto, alto, médio, baixo).
- **maint** (v-high, high, med, low): preço de manutenção (muito alto, alto, médio, baixo).
- **doors** (2,3,4,5-more): número de portas (2,3,4,5-mais).
- **persons** (2, 4, more): capacidade de pessoas no carro (2, 4, mais).
- **lug_boot** (small, med, big): tamanho do bagageiro (pequeno, médio, grande).
- **safety** (low, med, high): segurança estimada do carro (baixa, média, alta).
- **acc** (unacc, acc, good, v-good): aceitabilidade (inaceitável, aceitável, bom, muito bom).

Há uma amostra dos objetos desse dataset no capítulo do livro que usamos esta semana: [Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações, pag.177](#).