



Grupo 1: Predição de renda anual

Problema

O problema consiste no emprego de dados extraídos do censo americano, realizado em 1994, para inferir automaticamente se um dado indivíduo possui rendimento anual superior a 50 mil dólares.

Base de dados

A base de dados contém 48.842 amostras, cada qual composta por 14 atributos. Cada amostra, representa as respostas fornecidas por uma pessoa para 14 perguntas do censo.

A última coluna de cada amostra representa a classe, indicando se a pessoa em questão teve renda anual superior a 50 mil dólares ou não.

Referências

Para mais informações, consulte <http://archive.ics.uci.edu/ml/datasets/Adult> e o artigo:

Ron Kohavi, “Scaling Up the Accuracy of Naive-Bayes Classifiers: a Decision-Tree Hybrid”, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996.