

Trabalho 1

Faça uma análise de dois datasets, realizando o pré-processamento dos dados.

O primeiro dataset é aquele criado pelo professor com o objetivo didático, denominado “dadosAulav1”. Nesses dados, os alunos terão que converter todos os atributos em valores numéricos. Para os atributos nominais, será necessário utilizar o método de codificação de variáveis dummy, enquanto os atributos ordinais deverão ser mapeados em valores numéricos que representam cada grandeza. Por fim, os valores deverão ser normalizados em um intervalo de 0 a 1. Em relação aos dados, a informação sobre a "Doença" que estiver vazia pode ser considerada como “Saudável”.

O segundo dataset é referente ao Titanic, disponibilizado pelo professor juntamente com este documento. Esta análise pode ser realizada da melhor forma possível, não sendo necessário normalizar os dados ou converter todos os valores em numéricos.

O trabalho pode ser realizado em duplas. Envie os dados pré-processado e os comandos usados para essa filtragem com uma explicação do que o foi pensado quando executado cada comando. O nome dos arquivos terá que ser o nome dos alunos.