

Universidade do Minho

Licenciatura em Engenharia Informática

Aprendizagem e Decisão Inteligentes

3ºAno, 2ºSemestre

Ano letivo 2023/2024

Conceção de modelos de aprendizagem e decisão

ANEXO 1

João Magalhães, A100740

Jorge Rodrigues, A101758

Rodrigo Gomes, A100555

Simão Costa, A95176

1 Tarefa Dataset Atribuído

1.1 Estudo de dados

1.1.1 Age (idade), Year_of_birth, Month_of_birth, Day_of_birth (data de nascimento)

Achamos interessante ver a distribuição da categoria pelas diferentes idades

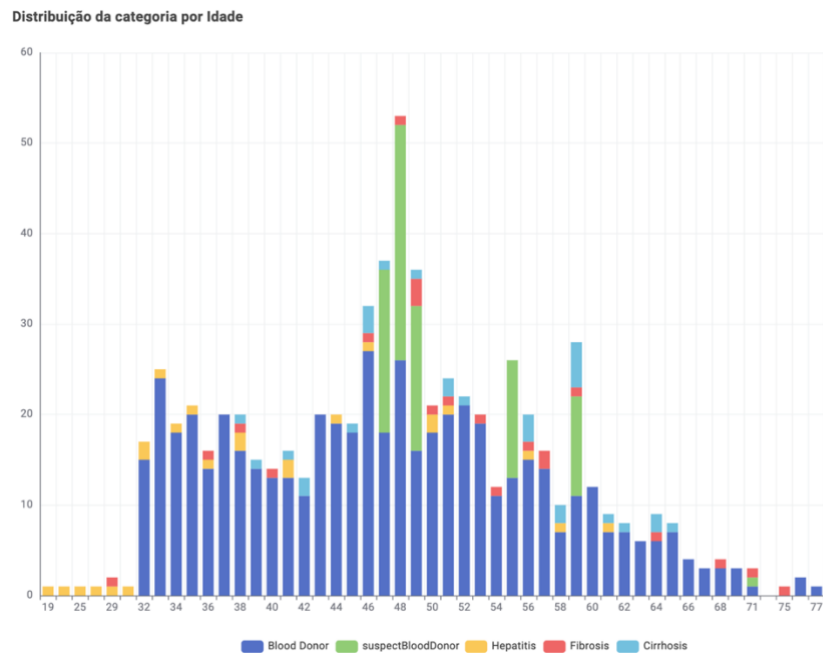


Figura 1 Distribuição dos diagnósticos por idade

Não conseguimos concluir muito à cerca deste atributo, tirando o maior número de casos percentualmente na população mais jovem.

1.1.2 Sex (sexo)

Após a limpeza de dados, tentamos ver se existe alguma relação evidente entre o sexo e o atributo alvo.

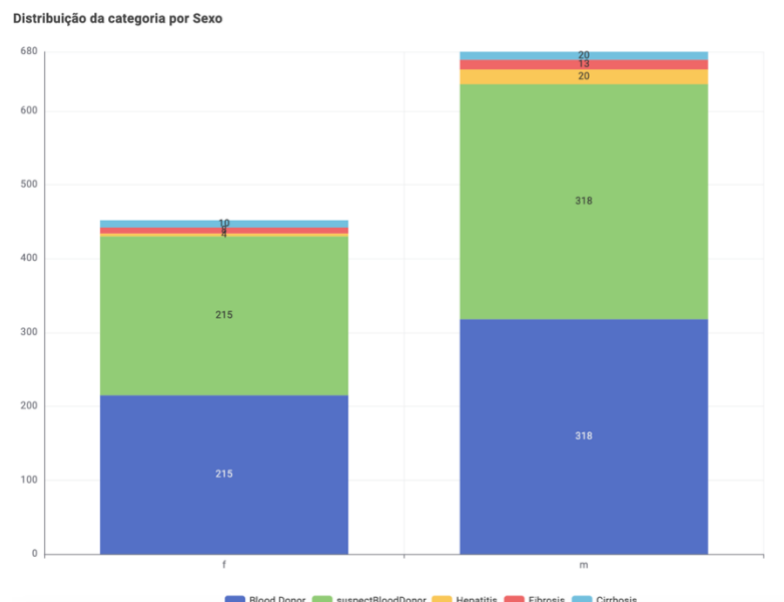


Figura 2 Distribuição de diagnósticos por Sexo

Com base no gráfico, tanto nos pacientes masculinos como nos femininos, o número de doadores de sangue é igual ao número de doador suspeito. No que toca a indivíduos infetados, em proporção, o sexo masculino possui mais doentes do que o sexo feminino. A proporção de infetados face ao total de pacientes femininos é de 4.84%, enquanto, no sexo masculino o mesmo ratio é de 7.5%. Deste modo, o género é um fator que consideramos que valha a pena considerar no diagnóstico de uma paciente.