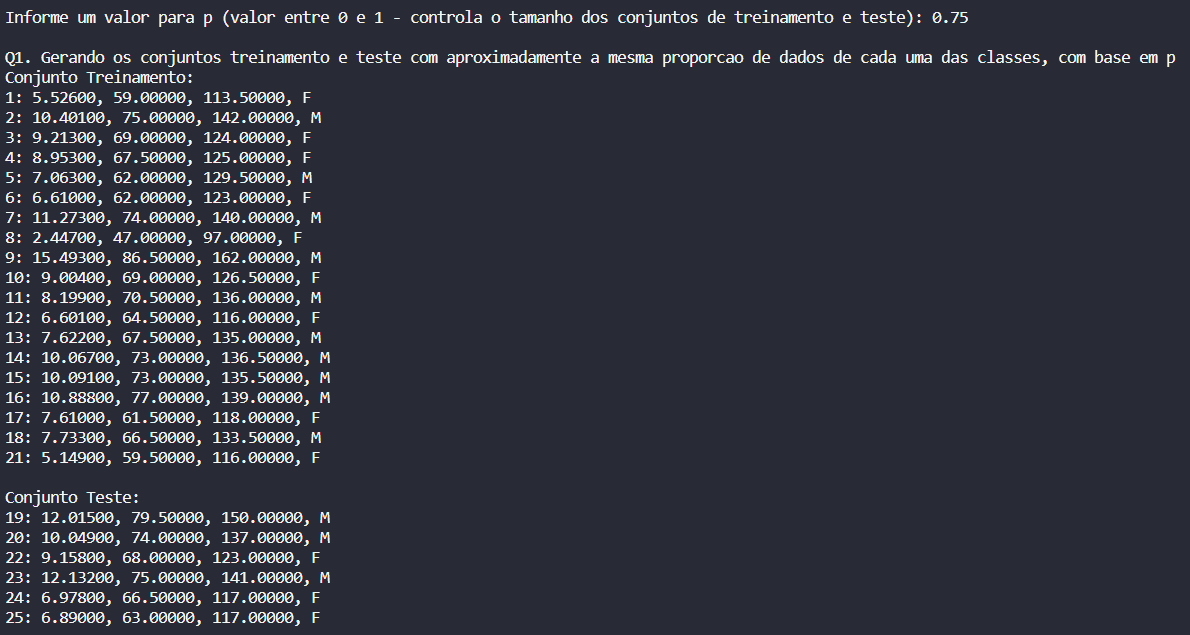
PRÁTICA 2

Larissa Magalhães Pereira

13747904

1. O programa faz a leitura do arquivo que contém todos os dados, armazenando as informações dos sexos dos lagartos em um vetor sexo[]. Posteriormente, ele pede ao usuário para informar um valor p (um número entre 0 e 1 que controla o tamanho dos conjuntos de treinamento e teste) e, com base nesse valor, separa os dados em conjuntos de treinamento e teste seguindo uma proporção igual entre as classes (feminino e masculino). Por exemplo, com p = 0.75 e sendo 13 machos e 12 fêmeas, temos que no conjunto treinamento terão 0.75\*13 10 machos e 0.75\*12 = 9 fêmeas, com proporção 10/9 = 1,11. Já no conjunto teste, terão 0.25\*13 3 machos e 0.25\*12 = 3 fêmeas, com proporção 3/3 = 1.

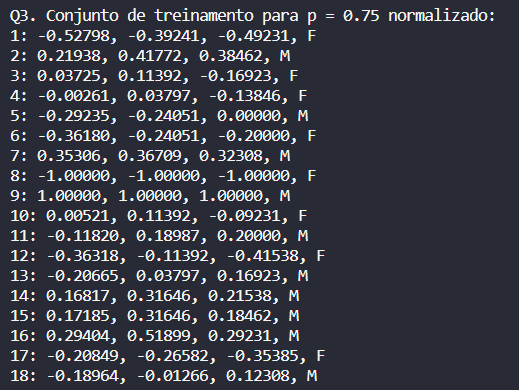


1. A fórmula usada para normalizar os valores dos conjuntos de treinamento foi:

* v = valor a se normalizar
* min = valor mínimo da coluna
* max = valor máximo da coluna

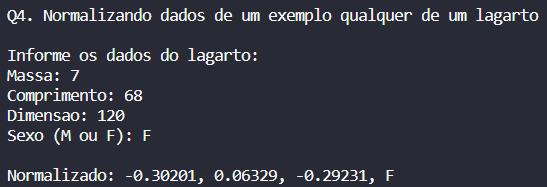
Vamos supor que se v = min, o resultado é 0. E, se v = max, o resultado é 1. Com base nisso, podemos deduzir a equação . Porém, queremos chegar em -1 para o valor mínimo e 1 para o valor máximo, assim, basta multiplicarmos essa equação por 2 e depois subtrair 1.

1. Normalizando o conjunto de treinamento para p = 0.75, temos:



O programa percorre 75% dos dados totais (conjunto de treinamento) por meio dos vetores massa[], comprimento[] e dimensao[], normalizando cada um de seus valores a partir da fórmula , sendo o max e o min respectivos a cada vetor/coluna.

1. Normalizar para um exemplo qualquer (com base no p dado anteriormente):



O programa pede o usuário para informar os dados (massa, comprimento, dimensão e sexo) do lagarto e aplica a fórmula ((valor - min)/(max - min)\*2-1) para cada variável, com base no conjunto de treinamento obtido anteriormente (em p e nos máximos e mínimos), obtendo a normalização dos dados.