**HEART DISEASE Dataset**

The problem scenario for this dataset is to identify whether a patient has heart disease or not.

This is a classification problem, using anonymised data from real patients, which has been made available freely online (however, note the dataset has been modified slightly for this unit). As it is part of your assignment to describe the dataset, only the following high-level information is given. You are encouraged to do some desk research to

The data consist of health records of patients that are classified according to the presence or absence of coronary artery disease; integer value from 0 (absence) to 4 based on the severity of the disease (in the final attribute).

***The problem should be solved as a multi-class problem, though you may create a binary version of the dataset for comparison if you wish (combining classes 1-4).***

While you are not expected to understand the attributes in terms of their medical meanings they are described as follows:

1. Age
2. Gender (1 = male; 0 = female)
3. Chest pain type
   * 1 = Typical angina
   * 2 = Atypical angina
   * 3 = Non-anginal pain
   * 4 = Asymptomatic
4. Resting blood pressure (in mm Hg on admission to the hospital)
5. Serum cholesterol in mg/dl
6. Fasting blood sugar > 120 mg/dl 🡪 1 = true, 0 = false
7. Resting electrocardiographic results
   * 0 = normal
   * 1 = having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)
   * 2 = showing probable or definite left ventricular hypertrophy by Estes' criteria
8. Maximum heart rate achieved
9. Exercise induced angina (1 = yes; 0 = no)
10. ST depression induced by exercise relative to rest
11. The slope of the peak exercise ST segment
    * Unsloping
    * Flat
    * Downsloping
12. The number of major vessels (0-3) coloured by fluoroscopy
13. Thalassemia (thal), a blood disorder: 3 = normal; 6 = fixed defect; 7 = reversible defect

Note that this dataset contains missing data, which you would need to handle as part of the assignment.

This is a modified version of a dataset available online:

<https://archive-beta.ics.uci.edu/ml/datasets/heart+disease>