**Lab 1**

The purpose of this lab is to get you up and running with Python version 3 and start playing with the search techniques.

**Setup**

Please ensure you have the following installed:

* Your preferred version of Python (I have been using 3.11).
* [Anaconda](https://docs.anaconda.com/free/anaconda/install/) is a good toolset for Python development. You can however just use a plain [Python](https://www.python.org/downloads/) install.
* We will use a virtual environment within Python to keep things clean
  + Create a virtual environment called “aiinthewild”:
    - python -m venv aiinthewild (this creates a directory called “aiinthewild”)
  + To activate it, run:
    - aiinthewild\scripts\activate
      * Now any python modules you install will get put in the new directory
  + If you want to deactivate it, run:
    - aiinthewild\scripts\deactivate

**Introduction**

Once you have Python ready to go, spend some time familiarising yourself with the fundamentals of python (check out the resources slides) then proceed with this exercise.

**Exercise**

Run through the Python K-Nearest Neighbours tutorial at the following address:

<https://realpython.com/knn-python/#use-knn-to-predict-the-age-of-sea-slugs>

The abalone dataset is on Blackboard for you already (abalone.data). Open it in a text editor to see what’s in it.

You should create your own abalone-knn.py file and cut and paste into it as you go through the tutorial.

You can **skip** the section “**A Step-by-Step kNN From Scratch in Python**” and jump down to “**Fit kNN in Python Using scikit-learn**”. We will use the powerful Scikit-learn library for this instead of writing the algorithm from scratch.

Note: It is NOT imperative that you understand every bit of the code. The idea is to clearly understand what it is doing and why, and then to be able to interpret the output.

----->

**To Do**

Repeat the process above using the Spambase dataset. The dataset is available on Blackboard and also described here: <https://archive.ics.uci.edu/dataset/94/spambase>

Adjust the code as you see fit so that a meaningful K-NN model of the data is created.

Submit your solution to Blackboard.