

61011 Lab - Week 1

David Wong and Christopher Yau

October 2020

1 Introduction

There are no in-lab assessments on this course.

This is a strategic decision so that you can focus on learning (and interfacing with the helpers in the synchronous lab drop-in sessions) without worrying about imminent deadlines. To help you self-assess your progress, we have provided a number of targets, at different levels of difficulty, for each week of the course.

2 Setup

2.1 Step 1: Install the virtual machine image (or other Python infrastructure)

All labs will be run in Python, as it is the main language used by data scientists and ML experts used worldwide. We will be using an Integrated Development Environment called Spyder to write and debug code.

To ensure consistency in programming language and library versions, we suggest you download the computer science virtual machine image (https://wiki.cs.manchester.ac.uk/index.php/CSImage_VM). Be aware that the setup requires around 50Gb of available disk space.

If you choose NOT/are unable to install the virtual machine, and you do not already have a development environment for Python, then install the Anaconda Python distribution platform from anaconda.com. If you already have experience in other IDEs (e.g. pycharm), then feel free to continue using it. Please note that NOT installing the virtual machine, may limit the amount of help that we can provide in the live lab help sessions (as we cannot be certain if problems might be due to the environment, rather than directly because of an error in the code).

2.2 Step 2: Check that you can access the live lab session

The live lab session will be run from 0900-1100 on every Wednesday of the course (starting 28/10, ending 25/11). The link to the session is provided in the course content on Blackboard. Check NOW that you can access the session.

During (or before) the live Wednesday lab sessions, you may request help by entering your name on the following spreadsheet:

https://docs.google.com/spreadsheets/d/1Ux0iZjqvF1MPciMzm1w_o5DviDSpTyg1PkiPKdtRBfA/edit?usp=sharing.

When you join the weekly blackboard collaborate session, a teaching assistant will move you to a private chat room, where you can discuss your problem and share your screen. We will acknowledge your request when we see it on the spreadsheet - please be patient, as there are a limited number of teaching assistants.

For all future weeks, you can request teaching assistant help *before* the session at sli.do using the code #92676 (this code will change each week). Please enter your full name, and 1 sentence description of your problem, and we will prioritise you at the start of the live lab session.

2.3 Step 3: Begin the Week 1 activities

By the end of this week (Friday) you should have completed all of the targets. From Monday, stop working on the Week 1 lab, and proceed to the Week 2 lab, even if you have not finished.

Level 0

- If you don't know it already, learn Python. Try www.learnpython.org
- Write a loop to sum the even number from 1 to 100.
- Sum the even numbers from 1 to 100 without a loop
- Write a function to return the sum of its arguments, and call it

Level 1

- Work through the tutorial at <http://scikit-learn.org/stable/tutorial/basic/tutorial.html>
- Use SKlearn to load a dataset, and display it on the screen.
- Use SKlearn to build a logistic regression model, then train it. You will learn what this model does next week
- Plot the testing performance of a logistic regression for different learning rates.