Methods for Data Science - MATH60026/70026
Coursework 2 Part 2 (CW2 Part 2) - instructions for students

Part 2 of Coursework 2 will take place on **21/03/2025, 9.30-12 am** in several Imperial College computer rooms (you should have received from the Maths Exams Office instructions about the room to which you have been assigned).

For this test, you will use a College computer and the Safe Exam Workspace (SEW), about which you have seen (and received a printed copy of) detailed instructions during the SEW familiarisation session. You will receive an updated printed copy of these instructions on the day of the test.

The session will proceed as follows:

- At 9.30am, you will start to set up SEW, following part 1 and 2 of the instructions about SEW under the guidance of the representatives from the Faculty of Natural Sciences Education Technology (EdTech) team that will be present in the room.

- Once everybody has successfully logged in and set up SEW, an invigilator will announce the start of the test. The coursework tasks are contained in the notebook, called 'CID_Coursework2_Part2.ipynb', that you will find following part 3 of the instructions about SEW. You will have 10 minutes to open this notebook, read the tasks, load the dataset (by running the first cells), and open some of the lecture notes and slides and notebooks available in SEW, but you *cannot* start solving the tasks.

- At the end of the 10 minutes for set up, you will work on completing the tasks in the notebook 'CID_Coursework2_Part2.ipynb' (see instructions below). The time allowed for this is 120 minutes.

- Finally, you will have 20 minutes for the submission of your work, following part 4 of the instructions about SEW. You will submit ***only*** your completed notebook.

Permitted materials:

You may consult:

1) the websites, whitelisted on SEW, with Python documentation (you see a list of them in the instructions about SEW);

2) all lecture notes, slides and solved notebooks for each week and other course materials made available within SEW;

3) your own hand-written notes contained in a A4-sized sheet of paper (single or double-sided).

Finally, during the test you can write your notes on a blank A4-sized sheet of paper that an invigilator will give you before the start of the coursework.

Very important: to access all the course materials (point 2 above), you will be able to connect to the course's page on Blackboard bb.imperial.ac.uk. This is also from where you will download the dataset for Coursework 2 Part 2 (`gene_expression_transcriptomic_data.csv` inside 'Coursework/Coursework 2/Data'). When you download notebooks and data from Blackboard, they should be automatically saved in the same folder 'Documents' on SEW. If you have trouble with this, ask an invigilator or the EdTech representative for help.

During the test, you are not allowed to:

- Use laptops, mobile phones, tablets and any device that can connect to the internet;
- communicate with anyone other than an invigilator, or an EdTech representative. This includes, but is not limited to:
    a. sending or receiving emails, or instant messages on any platform, and
    b. posting on any web forum.

And you must not attempt to:

- employ any conversational AI tools such as, but not limited to, ChatGPT, Microsoft Bing, or GitHub Copilot;
- Copy and paste code from online resources that are not whitelisted.

Using unauthorised materials or attempting to communicate with anyone other than an invigilator, or an EdTech representative, are exam offences and will be dealt with under the academic misconduct policy and procedures.

Note: make sure you bring the device (like your phone, or iPad) that you use for the authentication step when accessing your Imperial account in SEW. You will be given a folder by an invigilator where you can store your mobile phone during the test. Please make sure your phone is **on silent mode**, not just vibrate.

General instructions for completing the tasks

In CW2 Part 2, you will analyse the same dataset as in Part 1 using tools and algorithms introduced in the lectures, which you have also studied in detail through the weekly Python notebooks containing the computational tasks. You will solve the tasks in this coursework using Python.

You are allowed to use: Python code that you have developed in your coding tasks, including all the following Python packages that have been used there and that you will find installed in SEW:

- `numpy, pandas, matplotlib, seaborn;`
- `collections, typing, copy, heapq, tqdm;`
- `Torch`, but only to build, train and evaluate the performance of CNNs.

You are *not* allowed to use any other Python package.