CSSci

Code Portfolio and Explainer Session

Individual Assignment DE

Version 1.0 (20 January 2025)

The DE assignment consists of two graded components:

1) A portfolio of Python notebooks.
2) An oral Explainer Session on a selection of your portfolio.

Throughout the semester, you will receive coding notebooks. Your task is to complete these notebooks, which will include finding a dataset *not provided by us*, and to apply all methods learned in this notebook to your new dataset. You can pick amongst the datasets used for the Group Project V if necessary, though other datasets of your choosing are also possible.

For each notebook, your task is twofold:

1. To complete, and keep a well-annotated record of, the code and generated results from these assignments.
2. To familiarise yourself with the code and method enough such that if questioned on any part of the coding or modelling choices, you are able to defend your code.

It is your responsibility to ensure you complete the exercises properly. Periodically throughout the semester, you will have the opportunity to engage in a mock oral defence, where TAs / CLs will test your ability to defend coding and modelling choices informally, and give you feedback on how to better explain yourself verbally.

The portfolio is evaluated according to clarity and execution criteria, which are covered below. It is necessary for your code to run smoothly without hiccups in order to pass the individual DE assignment – the TAs in charge of replicating your code must be able to run it top-to-bottom on their own machine, and will be in touch with you if they cannot do this. You will have one chance to correct the code ***to ensure replicability***; to do so you must schedule an appointment with a DE TA on Friday, 4 April between 10:00AM and 17:00PM.

Unlike previous Explainer Sessions, you may not nominate or choose which notebooks you will defend. ***CSSci staff will chose between all submitted notebooks***. The defence will be conducted by a CSSci Core Lecturer and will be audio recorded; a TA or other staff member may also be present. The session will last around 15 minutes. Your job will be to:

1. Answer a series of questions related to your code, and/or
2. Provide accurate justifications for the chosen methodology's compatibility with your data, and/or
3. Provide accurate justifications for the chosen methodology's compatibility with the goals of the analysis, and/or
4. Comment on the results of your DE effort.

Deadlines

Monday, February 24th, 2025 10:00 - 13:00	**Practice Explainer Session 1** Optional but strongly encouraged.
Monday, March 10th, 2025 10:00 - 13:00	**Practice Explainer Session 2** Optional but strongly encouraged.
Monday, March 24th, 2025 10:00 - 13:00	**Practice Explainer Session 3** Optional but strongly encouraged.
Monday, March 31st, 2025 17:00	**Final code submission** Final submission via CodeGrade on Canvas.
Monday, April 7th and 8th, 2025 Time by appointment.	**Code Explainer Session** Details provided by your Core Lecturer.

If you have an unavoidable conflict on the days of the Code Explainer Session, notify your Core Lecturer as soon as possible. We may find an alternative time for your Explainer Session. If no alternative time can be agreed, you may be allowed to participate at the time of the repair opportunity as your first attempt.

Materials Provided to Students

Students receive a number of iPython notebooks which they need to complete.

Assessment Criteria

Coding Portfolio

- Completeness and Accuracy:
  The extent to which the code fulfils the given assignments and correctly applies the demonstrated methods.
- Documentation and Annotation:
  Quality of comments and documentation within the code. This includes clear explanations of the logic, purpose of different code sections, and documentation of the methods and data sources used.
- Innovation and Creativity:
  Assessment of the uniqueness and creativity in approaching the problem and data. This could involve innovative use of algorithms, unique data manipulation, or creative problem-solving strategies.
- Replicability and Robustness:
  The ease with which the code can be replicated and run on a different machine. This includes the absence of hard-coded paths or system-specific dependencies, and overall robustness of the code against different datasets.

Oral Explainer Session

- Mastery of Code and Methods:
  Ability to clearly explain the code, its structure, and the underlying methodology. This includes the rationale behind specific coding and modelling choices.
- Problem-Solving Skills:
  Ability to answer questions and defend coding choices effectively, showcasing problem-solving skills and adaptability in thinking.
- Communication Skills:
  Clarity, coherence, and effectiveness of verbal communication. This includes the ability to explain complex technical concepts in an understandable manner.

Repair Opportunity

To qualify for the repair opportunity, you must have submitted serious attempts for all notebooks as well as having participated in the Explainer Session.

Depending on your submission you will be asked to either:

- Resubmit (part of) your coding portfolio AND participate in the Explainer Session resit.
- Participate in the Explainer Session resit.

Learning Objectives

Students are able to:

- Demonstrate a thorough understanding of machine learning and other digital techniques for prediction.
- Understand how social biases manifest in digital data and predictive algorithms.
- Detect social biases in digital data and predictive algorithms.
- Implement appropriate social bias mitigation techniques.