Assessment: Machine Learning and Statistics

ian.mcloughlin@atu.ie

Winter 22/23

These are the instructions for the assessment of Machine Learning and Statistics in Winter 22/23. These cover the full 100% of the assessment for this module.

Submission

- The deadline for submission is January 7th, 2023.
- Your whole submission must be in a single GitHub repository.
- Use the form on the Moodle page to submit your repository.
- All you need to do is submit the repository URL.
- You should submit the URL as soon as possible.
- Commits in GitHub on or before the deadline will be considered.[1]

[1] Once you have submitted your URL, you do not need to do anything other than commit to your repository and push the changes to GitHub.

What to submit

This assessment has three overlapping components, as follows.

Presentation of GitHub repository (20%):

- Create a GitHub repository of a standard presentable in job interviews.
- Include an informative and concise README.
- Organize your repository — no unnecessary files or clutter.
- Work regularly, adding regular reasonably-sized commits.

Statistics Notebook (40%):

- Include all your practical work in your JupyterLite instance.
- The lecture notes suggest concepts you should explore[2].
- You should complete these in notebooks each week[3].
- Plots explaining main concepts.
- Code demonstrating key algorithms.
- Well-written explanations[4].

[2] The lecture notes themselves are presented in Jupyter notebooks.

[3] You don't need to have a notebook for each individual small topic. It is up to you to choose how you organize the notebooks. Aim for one notebook per topic.

[4] Pitch all your work at your classmates as the audience.

Machine Learning Notebook (40%):

- Create a repository for a custom JupyterLite instance.
- Add a GitHub action to publish the JupyterLite instance.
- Commits to the main branch should trigger the action.
- Detail all of this in your README.

Marking Scheme

Each component will be marked using the four categories below. To receive a good mark in a category, your submission needs to provide evidence of meeting each of the criteria listed under it[5].

[5] In line with ATU policy, the examiners' overall impression of the submission may affect individual marks in each category.

Research (25%)*:* evidence of research on topics; appropriate referencing; building on work of others; comparison to similar work.

Development (25%)*:* clear, concise, and correct code; appropriate tests; demonstrable knowledge of different approaches and algorithms; clean architecture.

Documentation (25%)*:* clear explanations of concepts in notebooks; concise comments in code and elsewhere; appropriate, standard README for a GitHub repository.

Consistency (25%)*:* tens of commits, each representing a reasonable amount of work; literature, documentation, and code evidencing work on the assessment; evidence of reviewing and refactoring.

Advice

- Students sometimes struggle with the freedom given in an open-style assessment.
- You must decide where and how to start, what is relevant content for your submission, how much is enough, and how to make the submission your own.
- This is by design — we assume you have a reasonable knowledge of programming and an ability to source your own information.
- Companies tell us they want graduates who can (within reason) take initiative, work independently, source information, and make design decisions without needing to ask for help.
- The point of this assessment is to demonstrate you can do that.
- You need a plan, you cannot just start coding straight away.

Policies

- You are bound by all ATU policies and any GMIT policies that have not yet been replaced by new ATU policies.
- Review the GMIT Quality Assurance Framework [6].
- Pay particular attention to the Policy on Plagiarism and the Code of Student Conduct.
- If you have any doubts about what is permissible, email me to ask[7].

[6] GMIT. Quality Assurance Framework. `https://www.gmit.ie/general/quality-assurance-framework`

[7] `ian.mcloughlin@atu.ie`