Final Project

1 Logistics

For the course project, you will implement a research idea related to the course material. The purpose of the final project is to give you some experience working on a piece of original research and writing up your results in a paper style format. You are expected to describe your research idea/application clearly in the project proposal, and relate it to existing work. You will present your ideas and results through an in-class presentation. Finally, you will document the project progress in the final report.

You **must form a group of two or three** to complete the project. Your report must clearly **list the contributions of each team member**, as well as **an estimation of the relative contribution of each member**. Each group member's mark will be adjusted based on how much they contributed. **Bonus points (10%) will be given to students who work in groups of 2, to account for the increased workload**.

- Proposal: Due Feb 20, 11:59pm

- Presentation: Due March 30 or April 6, in-class

- Final Report: Due April 13, 11:59pm

2 Writing format

All submissions must be in PDF format. You may include algorithm blocks, tables, and figures. The write-ups should be prepared in the Neurips Paper Format. You may find online editors such as Overleaf helpful for writing the reports.

Proposal: The project proposal is limited to two pages. It should roughly have the following sections:

- 1/4 page introduction

- 1/2 page related works

- 1/2 page method / algorithm

- 1/4 page abstract and reference

The point of the proposal is mainly for us to give you feedback and formulate a plan for the final report. The proposal will not be graded. We will set up project consultation appointments after we have collected all the project proposals. You will submit your proposal report through Quercus. **Note: Groups without proposal submissions cannot proceed with the final presentations and reports!**

Final report: You will expand out your project proposal to include experiments and comprehensive method sections. You are expected to discuss the experimental results in detail and

highlight any interesting findings. We recommend the final report to be four pages (excluding references).

You will submit your final report through Quercus. You must also submit the code necessary to reproduce your experiments, either with a github/colab link or as a compressed folder (.zip or .tgz).

3 Presentation Format

The slide presentation should be short, maximum 10 min. Presentations that are longer will be penalized, with a hard cutoff after 11 min. A good rule-of-thumb is one slide per minute (excluding title and reference slides). Effective presentations make use of large figures and minimal text.

The project does not need to be complete at this stage: it is more important to provide the audience with the necessary background, illustrate some of the key ideas, and show at least one preliminary result. All members of the group must be present for the presentation and have significant speaking time.

4 How to choose a project

The course projects should build on top of the course materials. You must apply one or two machine learning models to a biological or clinical problem in the biomedical sciences. You should compare your models with existing methods for solving the problem, and highlight the strengths and weaknesses of each approach. You may employ software and modelling techniques that were not presented in class, but we emphasize that the project is about **developing novel methods or applications**, not applying existing bioinformatics or health informatics tools.

You are welcome to do a project related to your research. In this case, your project proposal and final report must **both** clearly explain the relationship to your research, what work was already done prior to the course, and what work (if any) was done by people not on the project team. Our expectations will be higher in this case.

5 Grading scheme

The goal of the project is for you and your group to conduct original research. The presentation will be worth 30% of the total mark, with the report making the other 70%. The proposal will not be marked.

The presentation will be graded using the following criteria:

- **Term Presentation [30%]** Does the presentation provide suitable background and motivation? Are the figures/equations well explained? Is the presentation visually appealing and on time?

The final report will be graded according to the criteria of top machine learning conference submissions (NeurIPS):

- **Quality [25%]** Is the report technically sound? Are claims well-supported by theoretical analysis or experimental results? Is this a complete piece of work, or merely a position report? Are the authors careful (and honest) about evaluating both the strengths and weaknesses of the work? To get full mark in this category, you will need to include at least one of:

  - An algorithm box.
  - Equations describing your model.
  - A theorem or formally stated conjecture.

- **Clarity [15%]** Is the report clearly written? Is it well-organized? (If not, feel free to make suggestions to improve the manuscript.) Does it adequately inform the reader? Are the figures/tables properly labeled? (A superbly written report provides enough information for the expert reader to reproduce its results.)

- **Originality [15%]** Are the problems or approaches new? Is this a novel combination of familiar techniques? Is it clear how this work differs from previous contributions? Is related work adequately referenced?

- **Significance [5%]** Are the results important? Are other people (practitioners or researchers) likely to use these ideas or build on them? Does the report address a difficult problem in a better way than previous research? Does it advance the state of the art in a demonstrable way? Does it provide unique data, unique conclusions on existing data, or a unique theoretical or pragmatic approach?

- **Participation [10%] (We are counting on you!)** We will adopt a peer-review system through Quercus in which students will participate in reviewing the other classmates' reports. Reports will be made anonymously. We expect each student to review at least 2 reports. The participation score will be given based on the quality of the reviews by each student. You may find this guide helpful regarding how to write a good review. For participating in the review process, you will be awarded 10%.

Late submissions on Quercus, except in the case of an official Student Medical Certificate, will be accepted with 10% penalty every 24 hours from the deadline up-to three full days. So, you will get 0% on the assignment if it is submitted 4 days late.