

Project Information

General Information

The project overall corresponds to **50%** of your grade and is focused on the design of a creative computational solution or method to solve a real-world open-ended problem, considering application-specific constraints.

Topic: The topic of the proposal is up to you, and you can choose a topic that best fits your interests. If you do not have a clear idea for a project, one option would be to use problems used in machine learning competitions such as Kaggle or DREAM challenges. You can also reach out to me if you cannot find any project that interests you; however, it is recommended that students come up with their own topic/application for the project, since defining the project is the first step towards becoming an independent machine learning engineer and/or researcher.

Step 1 (Due Feb 9, 11:59 PM): You should first form groups of 4 students.

Step 2 (Due Feb 23, 11:59 PM): Proposal, 4% of grade

You should come up with a project proposal (see above for the topic) to solve an open-ended real-world problem using a deep learning solution that you will design. The goal of the proposal is to

- A1) Define the application you want to focus on
- A2) Formulate a problem and describe reasons that a solution is needed
- A3) Characterize the constraints* that are important for your application and for the problem
- A4) Identify the scope of the project
- A5) Your proposed solution that you want to design
- A6) How you plan to implement your solution
- A7) How you will be validating, testing and evaluating your solution
- A8) Describe alternative possible solutions and why your proposed solution could be superior
- A9) Feasibility assessment of project (among other things include the dataset you will use)
- A10) Risk analysis and what pitfalls may occur and how you will overcome them

* **Constraints:** in the past semesters, there has been some confusion as what we mean by constraints. Here, constraints refer to design constraints: considerations that your design need to satisfy in order to achieve the goals of the project. For example, a very simple constraint could be achieving high accuracy in a binary classification. Another example could be the format of input images that put a constraint on the architecture (requiring it to be translation equivariant). Constraints however do not refer to limit in the number of hours the team may have or difficulty in finding equipment. Those are related to feasibility assessment.

Your project proposal should be of ~2-pages long (<1000 words) containing all components below (capturing all 10 items above). Please use the bold-face items below as sub-titles within your proposal:

- **Background:** here, you incorporate items A1, A2, A3
- **Aims/Goals:** here, you incorporate items A3, A4, A5
- **Methodology:** here you incorporate items A5, A6, A7, A8
- **Feasibility and resources:** here you incorporate item A9
- **Risk analysis:** here you incorporate item A10
- **Role of group members:** explain how the work is divided

Your proposal must contain citations (list of references can be up to one extra page). Since the items A1-A10 will be used to grade your proposal, make sure that you address all 10 items. One easy approach would be to have subsections within each section (e.g., in background, you can have “Constraints: ...”). This also helps you to better organize your thoughts in the context of the items above.

Step 3: At this step, 1) you need to finalize the architectures that are relevant to your application (this may have to change multiple times throughout the project); 2) implement the proposed architectures (after hyperparameter tuning using a validation set, etc.); 3) implement or run the code of architectures or models that are related to your project, but act as baseline; 4) evaluate the performance of your architectures and compare them against other models; 5) start writing the report and prepare your presentation.

Step 4 (April 8, 10, 11): Presentation, 10%

On these days, you will present your work. (Note that April 11 is a Thursday but is scheduled similar to a Monday (since Monday April 1 is a holiday)).

The presentation allows you to improve your communication skills and learn from your peers. An updated version of the items in Step 1 should be used to guide your presentation:

- B1) Define the application you focused on
- B2) Formulation of the problem and reasons that a solution was needed
- B3) Characterization of the constraints that were important to consider
- B4) Description of your designed solution
- B5) Important (but summarized) implementation factors
- B6) Datasets and resources used
- B7) Evaluation results and comparison with alternative solutions**
- B8) How did you manage to satisfy constraints of B3 and what went wrong

**** Alternative solutions:** Make sure in your report and presentation, you report comparison with alternative models, including but not limited to state-of-the-art methods for a similar or identical task; (you need to run those models on your data and you cannot borrow results directly from the original papers).

I suggest incorporating the following items in each part of your presentation as follows:

Introduction/Motivation: B1

Problem Definition: B2, B3

Methodology: B4, B5, B6

Results/Evaluations: B7

Conclusions (what was found in this study and where to go next): B8

The presentation should be based on your project and a **rubric** will be provided to all students and will be used for grading by the instructor and the student audience. Generally speaking, it is recommended that for a X minute presentation you have less than X slides. Also, do not forget to have slide numbers. Also, take advantage of figures; avoid having busy slides with too many words. Use bullet points instead of full sentences if possible. Avoid having sentences or bullet points that go beyond 2 lines, etc. I will provide you with a **slide deck** as a sample of a 40-50 minute presentations (see the next section for list of different samples). Your presentation time depends on the number of formed groups.

I also have found this youtube video that provides some good tips for making good slides:

<https://youtu.be/Hp7ld3Yb9XQ>

Step 5 (Due April 12, 11:59 PM): Project Report, 30% of grade

IMPORTANT: Due to the date that I need to submit the final grades, I cannot extend the deadline for submission of the final project report. Particularly, since there needs to be enough time for peer-evaluation.

In this step, you would need to write your report and submit it by the due date. I recommend following the format and guidelines of Oxford Bioinformatics (“https://academic.oup.com/bioinformatics/pages/instructions_for_authors”) or alternatively NeurIPS (“<https://nips.cc/Conferences/2020/PaperInformation/StyleFiles>”). Either way, your main text of report should be ~5000-8000 words (excluding references, which is approximately 5-8 double-column pages). If you need more space to include more details, you can have a supplementary text document (free format) in which you explain details (without any word limit). You can also have supplementary figures and tables as needed but remember that your main document must be self-sufficient as the grade of the report is based on your main report. The supplementary material is only used if a specific detail is needed by the grader.

You should use the items below to guide your report writing (incorporate them in different sections of the report as appropriate):

C1) Define the application you focused on

C2) Formulation of the problem you considered

C3) Description of reasons that a solution was needed for the problem

C4) Characterization of constraints that were important for the application and problem

C5) Description of resources and datasets

C6) Detailed description of your designed solution (including implementation)

- C7) Detailed description of your evaluation approach
- C8) Description of alternative solutions
- C9) Description of results of your design's evaluation & how it achieves constraints in C4
- C10) Comparison of your design with alternative solutions in achieving the constraints in C4
- C11) Limitations and advantages of your designed solution

Your report **must** be organized under the following sections:

- 1- Abstract
- 2- Introduction: provide the background and motivation of your project. Incorporate items C1, C2 (at high-level), C3, C4
- 3- Methods: describe (in details) the methodology. Incorporate items C2 (more details), C5, C6, C7, C8
- 4- Results: describe (quantitatively) how your design achieves the constraints in C4 and compare to alternative methods. Incorporate items C9, C10 (quantitatively).
- 5- Discussion and Conclusions: discuss the results and conclusions. Explain how (and whether) your designed solution achieved the constraints in C4, and if not, why. Incorporate item C11.
- 6- References

Your reports will be graded by the instructor and by your peers. In fact, each group will be responsible to peer-review two other project reports and the peer-review report is 6% of your total grade in the course. A **rubric** will be provided to you in advance.

Step 6 (Due April 19, 11:59 PM): Each group's peer-review reports (2 per group) are due on this day. A rubric and instructions will be provided to assist you in the peer-review process.

Important Deadlines

Deadline to Register your group: Feb 9, 2024

Project proposal deadline: Feb 23, 2024

Project presentations: April 8, 10, 11

Deadline to submit your final project report: April 12, 2024 (no extension is possible)

Deadline to submit your peer-review: April 19, 2024

Samples and References to help you with your project:

1- Sample **Slides**: A set of slides for a 40-50 minute presentation is uploaded. Please keep mind that your project presentation will be much shorter.

2- Presentation grading **rubric**: A rubric will be shared with you to know how you will be graded (by the instructor as well as your peers) for your project presentation. Pay careful attention.

3- Sample **report**: Two examples manuscripts from my group:

- <https://doi.org/10.1093/bioinformatics/btac383>

- <https://doi.org/10.1093/bioinformatics/btac429>

4- **Proposal** guidelines: The actual format and required components of proposal is discussed above. However, the link below provides some good suggestions and recommendations:
<https://www.uh.edu/~lsong5/documents/A%20sample%20proposal%20with%20comment.pdf>

5- Report grading **rubric**: A rubric will be shared with you to grade your peers' project report. The same rubric will be used by instructor to grade your project report.