

Project report requirements

Project reports should be written in the style of a *Bioinformatics* journal article. Word and LaTeX templates are available [here](#). You do not need to follow the [article guidelines](#) imposed by the journal. Reports will be graded based on content rather than length, and 4-6 pages is the permitted length range. A well-written 4 page paper is preferable to a rambling 6 page paper.

Strive to identify a specific conclusion and use your results to build support for your conclusion. Positive or negative results are equally acceptable; it is fine if your data do not support the initial hypothesis you proposed. Avoid reporting inconclusive results if possible. If you must, use the Discussion to explain why the results are inconclusive and propose additional tests you could hypothetically perform to arrive at a definitive conclusion.

Your report should be written with class members as the intended audience and can assume that the reader is familiar with the papers we have read in class. Specific sections include:

Title

Abstract: Provide a one paragraph overview of the problem you studied and your results.

Introduction: Introduce and motivate the problem you selected by providing relevant context and related computational work. State the hypothesis of your project clearly and preview whether your conclusions support or refute this hypothesis.

Methods: Describe your approach and data analysis. Include sources of data, third party methods you used, and new algorithms you implemented as applicable. Include details that are necessary for reproducing your results such as parameters used and how you selected them.

Results: Describe the outcome of your analyses. When you describe quantitative results, figures, or tables be sure to explain how they provide evidence for the conclusions you have drawn. Place any biological predictions in the context of biological literature.

Discussion: Summarize **and interpret** the results of your study. Interpretation of the results is critical. This section can include potential future work.

Figures and Tables: Reference all figures and tables in the text and provide informative captions. Be sure that figure annotations are large enough.

References: Use an author-date citation style and provide citations to related work, data sources, software, and relevant biological literature. You may include URLs as references, footnotes, or inline text.

Supplement: Include all code used in your data analysis as supplementary files along with a README file describing the files. These will be kept private. Unlike the homework submissions, your code does not need to be executable on the biostat server because it can utilize arbitrary dependencies. Rather, the code will be used to confirm that you have a strategy for reliably reproducing your results on your own hardware. You do not have to submit third party dependencies or software that you used.

Rubric

The project report will be graded as follows:

- **Formatting (10%):** Are all required sections present? Were the length limitations followed? Are the references complete and correct? Are there excessive typos?
- **Methods (20%):** Are new algorithms described in sufficient detail? Is sufficient information provided to reproduce the analysis (for example, algorithm parameters)? Is it clear where data were obtained and how they were pre-processed?
- **Results (20%):** Are the figures and tables legible and understandable? Were they selected to present a specific point? Do the selected evaluations provide appropriate data to determine the relative merits of different algorithms or justify the biological conclusions?
- **Conclusions (20%):** How were the results interpreted? Is there an explanation of why one method outperforms another or future work that could be conducted to assess this more deeply? What is the biological meaning of the computational results? What insights were gained from your analyses?
- **Overall clarity and scientific rigor (20%):** Were the methods and project strategy appropriate for the biological problem? Have alternative explanations of the results been considered? Is there a logical flow to the paper?
- **Supplementary code (10%):** Have supplementary files been provided? Are they convincing that the analysis can be reproduced?

The descriptions above are not comprehensive or equally applicable to all projects but serve to illustrate what features will be assessed. Plagiarism - ***including using text, images, or code without attribution*** - and cheating will not be tolerated and will be dealt with in accordance with the [Academic Misconduct Process](#).

Late days

There will be no late days allowed on the project report. Grading will begin immediately on Thursday May 11.

Submission

Submit your project report in PDF format to the following directory on the biostat server:

```
/u/medinfo/handin/bmi776/project_report/<USERNAME>
```

where <USERNAME> is your biostat user name.

Submit your associated supplementary files to the following directory on the biostat server:

```
/u/medinfo/handin/bmi776/project_supplement/<USERNAME>
```