PubH 7440: Final Project Instructions & Guidelines

Spring 2022

The final projects are to perform a Bayesian data analysis project on a realistic task. Through the final projects, you are expected to

1. *demonstrate your knowledge of Bayesian analysis techniques,*
2. *exhibit the advantage of Bayesian methods in certain data analyses.*

The projects are to be done individually. You may choose your own topic for the final project. The sequence of stages in the project process are

**Project Topic due Friday Apr. 1 via email**

You should decide a topic for your final project. There are two directions for the projects that you can consider:

1. The project may be very practical in terms of **applying existing Bayesian methods** you learned in or outside of class to a real problem.

2. The project may involve **designing or adapting existing algorithms** to a novel scientific or statistical problem.

The project can be related to your research area (if you have one). Do not submit anything you have completed prior to attending the course.

**Project Proposal due Tuesday Apr. 12 via email**

To judge the suitability of your choice of a project, you should produce a 1-2 page Project Proposal first that describes your idea for the project, and the work you intend to perform. This proposal should outline:

1. A high-level statement of the problem you intend to address

2. The data source(s) you intended to use

3. How you plan to obtain that data

4. The goals of your analysis

5. A description of the Bayesian analysis tools you plan to use

6. The products you plan to build, ideally including visualizations, and a report of outcomes.

The proposal will then be assigned to another student for peer review.

**Peer Review of Project Proposal due Thursday Apr. 14 in class**

You will be assigned to review a Project Proposal written by another student. The reviewer should write a report to provide feedbacks and constructive criticism for the reviewed proposal. You may also ask critical questions for the authors to address. I will spare some time in class so that the reviewer and reviewee pairs can meet and discuss over the review reports.

**Project Presentation Apr. 28 in class**

You need to prepare a PowerPoint presentation of around 20 min. The presentation should include

1. Problem statement: What scientific question you want to answer? What dataset you will analyze to answer the relevant question?

2. Methods you used and the rational for their use: May include data preparation/preprocessing, and then the Bayesian model(s) that you tried, performance measurement etc.

3. Results (may be preliminary): Any results you have to report so far. May also report unexpected challenges.

4. Lessons learned and/or plans to mitigate challenges.

**Final Report due Friday May 6 via email**

The project report is the formal description of your project. The report should be 5-10 pages in length (preferably using Tex). The structure is similar to the poster presentation but with full elaboration on what you did.

1. Problem statement and background

Give a clear and complete statement of the problem. Where does the dataset come from, what are its characteristics?

Include background material as appropriate: who cares about this problem, what impact it has, what implications better solutions might have.

Included a brief summary of any related work you know about.

Overview the Bayesian methods/ hierarchical models that you plan to use and justify your methods in terms of the problem statement.

Include measures (e.g. RSS, LPML, DIC, WAIC, LOO-CV) that you plan to use for model assessment/selection.

2. Methods

Describe the Bayesian methods you used in detail. Justify your choice of the hierarchical models, priors, and parameter specifications.

Describe the computational tools that you use for the Bayesian methods.

3. Results

Give a detailed summary of the results of your work. Present the data analysis results and the measures that you used for model assessment/selection.

Present your findings from the analysis in terms of your scientific questions.

Please use tables and figures to display your results whenever possible.

4. Discussion

Elaborate the lessons learned and/or plans to mitigate challenges.

**Grade Policy**

The grade of the final project will be based on the proposal (20%), proposal review (10%), presentation (30%), and final report (40%).