Final report requirements for BSC2022

Yanfei Kang   2022-06-01

Every student in this course (Bayesian Statistics and Computing) is required to complete a course project with a final report. You are expected to address some questions that interest you with the statistical computing methodologies we learn in this course. In this process, you will synthesize and rethink what we learn. I suppose it is a very good way of solidifying your understanding of statistical computing methods.

The report should at least consist of

- Introduction.

- Methodology.

- Discussion and conclusion.

Some ideas for reports

Basically, you choose your topic: anything we learn in this course, or their extensions. And you decide how to collect the data: you can either find some interesting data over the internet, or collect some data by web scraping.

You might have some ideas based on the following topics:

- Optimization: applying SGD to advanced regression/clustering/classification tasks.

- Bayesian inference: applying Bayesian inference & MCMC to solve regression/clustering/classification tasks.

- Applied linear algebra: SVD for image recognition (e.g., recognize whether a person wears a mask), or SVD for text mining.

Due Date

- July 6, 2022.

Report length

- Less than 10 pages (A4, 11pt, 1.5 line spacing).

Languages

- Computer language: **R** is preferred.

- Chinese and English are both acceptable.

Files for submission

Please submit a zip file that allows me to reproduce you results for marking. The main report should be in **pdf** format. Please do not include computer code in the pdf report.

Report grading guidelines

When grading your report, I will be looking for the following characteristics:

- Clarity: From your report, is it easy to understand what you did and the arguments you made? (20%)
- Interest: Did you tackle a challenging/interesting question? (20%)
- Relevancy: Did you use statistical computational techniques wisely to address the question you want to solve? (20%)
- Coding: Did you code up the main part by yourself, or just use some existing **R** packages? (20%)
- Comparison: Did you compare your methods with some benchmarks, and achieve comparable accuracy/efficiency? (20%)