Final Project

For your final projects, you will **work in a group** to implement computational/ machine learning algrithms **from scratch** using what we have talked about in class. You can either implement an algorithm we have discussed in the lecture (but without using the exactly same code provided) or any method you find interesting/ relevant to your research. After the implementation, you are also required to use Shiny to launch a dashboard to present your work.

Collaboration Guidelines:

1) You can either work in a group of two or three members.
2) If you work in a group of two, you will be expected to implement **at least** two algorithms. If you work in a group of three, you will be expected to implement **at least** three algorithms.
3) The algorithms can not be **completely different** from each other: functions for linear regression, logistic regression, and decision tree is the same project are not allowed. Please find a consistent topic and identify the connection among the algorithms. For example, you can discussion different modifications on boosting/ efficient matrix compression techniques/ acceleration of algorithms etc.
4) You are welcome to explore algorithms which are not mentioned in the lecture.
5) As a team, you are required to submit a report, a presentation slides, a shiny dashboard and other supporting codes/datasets.

Code Requirement:

1) You can choose your own coding language: R/ Python are preferred. Rcpp is optional.
2) The algorithm you are going to implement could be covered/uncovered in the lecture. But **you can not reproduce** exactly the same code from the slides.
3) The core part of your code should be **coded from scratch** (i.e don't use sklearn modules if you want to implement SVM), for preprocesssing and standarzing, you can use built-in libraries.
4) Once the algorithm is complete, **test and check** the correctness of your algorithm with data (synthetic data/ real data)
5) Your performance will be assessed based on the creativity, originality and thoroughness of your work.

Shiny Dashboard requirement:

Use your selected data to create a Shiny app or dashboard that allows users to interact, explore, and better understand the data. You may want to work with the Navigation Bar Page layout or Shiny Dashboard. Both options will neatly structure your app's user interface; however, these are not the only options.

Required app features:

1. An information section - what is the data, where did it come from, what is the purpose. Make it clear to users what data they are exploring.
2. It should have at least three different input/control widgets.
3. It should have at least three different output displays.
4. It should be well organized and aesthetically pleasing.

5. (optional) Feel free to incorporate at least one derivative shiny package such as shinythemes or shinyalert. There are many more, these are just two examples.

Check out Shiny's gallery for inspiration.

You will write up your analysis in a written report, and make an oral presentation. The presentation will be at most **15 minutes** in total.

Expectation for your presentation and report:

1. Give an introduction to the algorithm you want to implement. (If it has been dicussed in the lecture, please be concise)
2. Provide discussion on your motivation.
3. Describe the underlying theory and numerical implemention for the algorithm. Clearly state how your numerical implemention is been performed step by step.
4. Evaluate your work with synthetic data/real data and describe your result.
5. Discussion: Provide discussion of your algorithms. For exampe: comment on the efficiency of your algorithm, provide literature review of how others are implementing the same algorithm, potential improvement on your current work.
6. Speculate and generalize but use careful language. Say "It seems" or "appears" rather than "is" when it comes to speculative statements or models. For example, you might say "The residuals appear approximately normal" or "a linear model seems to fit well" but not "The residuals are normal".

7. In the written report, you can consider sections on introduction, algorithm and implementation, result, and discussion. Your report should be 5 pages or less of text, followed by as many figures or tables as you want.
8. In your oral presentation, start with an introduction explaining which algorithm you have implemented, then how you designed the numerical computation, and evaluate your code with sample data, evaluation with figures are preferred, and then conclude.
9. You have to turn in the written report, scripts and presentation slides by 12/8. But you can moderately modify your presentation slides before your oral presentation. (please make sure to upload the modified version to canvas!)