ICMF 324 Big Data and Financial Analytics

Trimester 2/2024

Guidelines for the Final Team Project report (15 marks)

One of our goals is for you to produce a piece of work you can use to demonstrate your data and programming skills.

Below are the step by step details and format that I ask of you to post your project in.

Not following these instructions exactly will result in a 10 percent reduction in the available points that your project can receive.

Due Date: Midnight, April 7, 2025.

If you are working in a group, each individual must post the group project on their own google classroom. So if Wasin and Wasinee are working together, then Wasin posts the project on his google classroom, Wasinee posts the project on her google classroom as well.

Objective:

The goal of the final project is to gain practical experience with a particular ML task or data analytics. Example tasks include topic Decision Tree, Regression with regularization, Logistic Regression with regularization. You are free to choose any task as long as it involves some kind of predictive or exploratory analysis. You are highly encouraged to choose areas that is interesting to you!

Overview:

At a high level, you are expected to: (1) select a particular machine learning task; (2) find or produce a dataset that can be used for experimentation; (3) design a program or use an existing toolkit to test one or more hypotheses; (4) do evaluation analysis; (5) report your findings. The goal is not for you to invent a new algorithm, but rather to apply existing techniques to a new application, or explore different feature representations on an existing task.

Guidelines:

Final report : Your final project report must be 25-30 pages. Your final report should address or have the following sections:

Introduction: What ML/data analytic application are you studying? Why is this an interesting problem? Why are your experiments interesting?

Approach: What exactly is your approach? What are its inner workings?

Experimental methodology: Describe your experiments and the metrics that you have used.

Results: What do you observe?

Discussion: From your results, you have to discuss the interesting

observations and to explain why something happened. If you found nothing interesting that is fine. But you must explain to me why you find nothing, Is it something to do with the data, methods, or anything else?

Conclusion: Overall, what do you do, what do you find, and why is that important?

Submit:

1. Final Report.
2. Python codes and Data set.

- Your Jupyter Notebook. This should be titled in the following way:

  lastname final project.ipynb

- A .pdf file of your Final report. This should be titled in the following way:

  lastname final project.pdf

Grading

Projects will be graded on their overall quality. This includes, but is not restricted to, these categories:

- Did you follow directions. Turn in data report on time, name files correctly, etc.
- Quality of the idea. Is the question clearly articulated? Is it interesting?
- Quality of the data. Does the data support the idea? Is it the best data for this question?
- Quality of the code. Is the code readable? Could someone else understand what you were doing and why?
- Degree of difficulty. Some ideas are harder than others to implement. As in Olympic diving, you get credit for taking on a challenge.
- Professional look. Does your project look professional? Are the graphs easy to understand? Are they clearly labeled?