

# MATH6010A – TP A

David Ardia

version 2023-08-27

## Important information

- This assignment is worth 20% of the final grade.
- This work must be done in teams of 3-4 people.
- The 10% oral interview will validate this work individually.
- I will only answer questions asked on the forum.
- Clearly document all your code.
- The total number of points is 100. Points per question are indicated in square brackets.
- You may submit your code in a language other than R (e.g., Python or MATLAB).
- Your files must be deposited no later than the deadline. 10% points will be retri ved for each hour exceeding the deadline.

## Objectives

Your objective is to create a realistic investment strategy based on several machine learning techniques seen during the course. The strategy is your choice: It can be a stock picking or market timing strategy, a geography/sector rotation strategy, currency trading, crypto trading, futures momentum/mean reversion trading, etc.

## Requirements

While you have total flexibility for the strategy creation, you must fulfill the following requirements:

- The data should be different than the one seen in class. Browse the web!
- The data dimensions should be at least 50 (features, assets, etc.). Be creative!
- The backtest period should be at least five years at the monthly frequency (you can use lower frequency features, but the backtest should be at least monthly).
- The machine learning techniques should include at least a regression and a classification technique.

## Deliverable

You must submit a report and the code to reproduce your results.

### 1. Report [50 points]

You must provide a three-page report (font Arial, size 11). The report must contain:

- a) A broad description of the trading strategy and its performance (1 page) [20 points]. You can use plots and tables for the performance metrics. Look at funds fact sheets!
- b) A detailed description of the machine learning strategies used (1 1/2 page) [20 points]. You can use maths to explain how you build signals.
- c) Comments about pros. and cons. of the approach (1/2 page) [10 points].

## 2. Code [50 points]

You must provide me with a zip file containing your code and data. The overall structure should be organized as a Rstudio project and contain the following:

- A `run_install_packages.R` file to install all required packages.
- A `run_strategy.R` file to run the backtest.
- A `data` folder with the raw data and a `README` file explaining from where they have been taken (with the steps)
- A `run_data.R` file to process the raw data if necessary or retrieve the data from the web.
- A `functions` folder with R files containing your functions.
- A `outputs` folder with all the tables and plots generated from `run_strategy.R`.

I will evaluate your work as follows:

- a) Project structure [10 points]
- b) Coding convention [10 points]
- c) Coding efficiency [10 points]
- d) Coding correctness [20 points]

The code should be documented and run without any issues.

### Bonus

- You get +5 points if you use RMarkdown.
- You get +5 points if you use data from Eikon Thomson Reuters or Compustat (available from the Laced).
- You get +5 points if you use data from OptionMetrics (available from the Laced).

### Some ideas

- Read carefully chapter *R programming* of the manual *Statistical methods for quantitative finance*.
- Use the signals to filter the universe.
- Use double sorting.
- Plug the signals into an efficient frontier setup.
- Plug the signals into risk-based portfolios.
- Use the “portfolio from sorts” approach.