Advanced Machine Learning in Finance, COMP0162, 2024/25

Cohort 2024/25. This assignment is worth 100% of the overall mark. It will need to be submitted to Moodle before the deadline of 25/03/2025.

Standard and non-standard calculators are permitted

Written report A single written report regarding an application of advanced machine learning in finance in pdf (maximum 10 pages) structured into:

- **Introduction.** The introduction describes the motivation and background of your project idea and includes the literature review with a selection (maximum of 10 articles) of relevant publications that are clearly linked with the project.

  The introduction should explain in plain terms the task under investigation, its relevance in finance or economics, and clarify why a machine learning model is suitable for the problem under investigation.

- **Methodology.** The methodology section should include a clear mathematical formulation of the machine learning model implemented, presenting the main formulas and equations that allow the reader to understand the full logic and functioning of the model and of the experiment.

  The main methodology should be based on a neural network model presented during the module. The section should include also a brief description of the null baseline model to compare with.

  The methodology section ends with a brief description of the data, it should explain where the data were obtained, describe the exploratory data analysis, explain and justify the data cleaning procedure, if any. It can include pseudocode and flowcharts.

- **Results.** The results section includes a detailed presentation and analysis of the results.

  Figures and tables should allow the reader to easily identify the key findings in the results.

  The results should compare the main machine learning methodology with a baseline model, e.g. a linear or logistic regression, an autoregressive model, preserving the main statistical features of the data.

  Results should include the statistical validation methods applied with a detailed explanation of the performance measures utilised and the error analysis.

- **Discussion.** In the discussion section, results are critically interpreted and connected with the main mathematical assumptions, and properties of the applied methodologies, outlining the advances, relevance for the financial task under consideration, limitations, and perspectives for future studies.
- **Bibliography.** It presents a clear list of the resources and papers referred to in the main text.

Coding, Editing, and Data

- LLM tools should be used only for editing and not for original writing.
- Students can use any programming language and any editing software for the report.
- The code will need to be uploaded as well (in your preferred format, e.g. ipynb) and separately from the main pdf.
- Students need to fill a Ethical Risk Identification (ERI) form regarding the data used for the project and upload it with the submission.

The final submission should be essentially three files, the main report, the code, and the ERI form. The assessment will be conducted on the main report.

Marking Marking will be based on the following criteria (a detailed rubric is shared on Moodle):

- Background information and introduction;
- Project structure;
- Clarity of presentation and explanations;
- Content and results;
- Consistency of language, code and mathematical notation;
- Critical interpretation of results and further challenges.

All projects will be thoroughly checked for plagiarism.