FE 691: Deep Learning in Finance
Final Project
Due Wednesday, March 1 11:59PM Eastern Time

Introduction

For your final project for the semester, you will work *individually* to use neural networks and deep learning to build a tool for financial analaysis (e.g., to do predictive analytics or portfolio optimization) utilizing financial data. Broadly your projects will consist of three inter-related topics:

1. accessing and discussing applicable financial data;

2. selecting an appropriate methods and hyperparameters; and

3. results and analysis.

   You will get the most out of the project if you interact with Professor Feinstein during this assignment, especially when planning a topic.

Project Components

Data Collection and Discussion

Download financial data from the database of your choice (e.g., Bloomberg, WRDS, *Yahoo Finance*, ...). You should spend time deciding on the appropriate data to analyze and whether it is sufficient for the method you will want to implement.

Machine Learning Methods

Given the data you have collected, choose and test machine learning method(s) to analyze your data. You can consider comparisons between other methods (e.g., linear regression or random forest) as baseline models. Give serious thought to your proposed methods as you will need to justify your choice.

Results and Analysis

Implement your methodology on your collected data in order to test the results. You may want to compare your chosen methodology to a simple baseline in order to determine performance. You should remark on whether your methodology appears suitable to answer the desired question; statistical analysis is *strongly* encouraged.

Report Details

You will submit your final write-up, which should include all of the information detailed below. This should be presented in roughly the order given, but your write-up need not have corresponding sections or bullet points. The write-up should be about 5-7 double-spaced pages, Times New Roman 12pt font. This does not include any appendices (of, e.g., your *Jupyter Notebook*) you may wish to include. Any external resources used should have clear citations and a reference page at the end of your work. This report **must** be submitted in pdf format; your code may be requested if not clear in the document so please keep that available.

1. **Overview** of the problem statement.

2. Detailed description of the **data collected** and why it is appropriate for the problem being considered. Mention any data cleaning if required.

3. Detailed description of the **machine learning method** and why it is appropriate for the problem being considered. If comparison to a *baseline* model is to be studied, provide the details of this methodology as well.

4. Describe the **results** obtained by your methodology on the data. Analyze these results to provide a recommendation.

5. **Next steps**: What do you recommend as a result of your analysis? Do you suggest attempting different algorithms or a larger test or more data? etc... What else could be done with the problem, but time did not permit?

Project Presentation

The final lecture of this course (Thursday, March 2) will be dedicated to your presentations. All of you will give **10-15 minute presentations**. An additional 1-2 minutes will be used for questions from Professor Feinstein as well as the audience. Slides must be uploaded to Canvas at least 30 minutes prior to class on Thursday, March 2.

If you have to attend remotely, you will need to upload a recording of your presentation to Canvas as well. This is solely in case of internet connection errors. All presentations are to be given *live* during the class time.