

## Stock price prediction using historical data and recurrent neural network

### 1. Project Description

This project studies the usage of historical data and recurrent neural network to predict a stock price or an index. Students can choose to use regression (prediction of a stock price or an index) or classification (prediction of the upward or downward trend of a stock price or an index).

### Data Generation

1. Goto <http://finance.yahoo.com>
2. Search one (or several, depending on the need for your model) of the stocks from a company (apple, amazon, microsoft, etc.) or one or several stock indices (S&P 500, Dow 30, Nasdaq, Russell 2000, Crude Oil, etc. )
3. Once the stock or index's quote page is shown, click the "Historic Data" tab and change "the time period" as needed. Note normally we prefer more data for a neural network.
4. Click "Apply" and then click "download" below the "Apply" button and a CSV data will be generated and downloaded to your computer.

note: You need to pick a period and decide how many data points to produce. Like the example using NYSE data in class or in homework 5, you need to rearrange the data into time series by lagging the data. You also need to split your data into the training set and the test set thus the validation can be performed.

### Recurrent Neural Network

You need to setup a RNN model and use the obtained data to predict stock price or its trend. Consider the following questions when you setup the RNN.

1. What's the overall dimension of the RNN?
2. What is the number of time steps? (Note: in the homework 5 example, we hardcoded the lags as 5; In this project, you need to show how to use a parameter to control the number of time steps; e.g. You only need to change one place when you switch from model with 5 lags to model with 10 lags.)
3. How many neurons are in the hidden layer?
4. Did you use the LSTM, if so what are the related parameters?
5. What is your choice of activation function? Why?

### Research Objective

Through this project, you will demonstrate:

- (1) You can clean and rearrange data in the form that a neural network can be applied on.
- (2) You can setup up a RNN model that is working.
- (3) You can further tune the model by modifying model parameters to finalize an optimized model.
- (4) The creativity based on accumulated knowledge and skills.

## 2. Project Specification

Your report should be produced using *Jupyter notebook*, by which you can use Markdown and code alternatively to explain what you have done and what you have found. The report should contain the following parts.

- (1) (2pts) general information: project title, course, section, date, name, department/program, etc.
- (2) (10pts) project description: describe how you obtain the data, the data, and what are your methods. Give an introduction to the stock(s) or index(indices) you have picked and explain why these are your interests. Describe your model. How many parameters does your model has, what are the predictors, what is the response?
- (3) (10pts) code and explanation of the code (use comments to explain the code as needed). Again, the number of lags should be controlled by a parameter.
- (4) (10pts) numerical results and discussion: evaluate the performance of your model. Adjust your model by modifying the parameters and try to optimize the model.
- (5) (2pts) concluding remarks: summarize the project, and raise questions if there are any.
- (6) (4pts) the report should be given in a neat, organized, readable, and professional way. You may use section, markdown, and other functions of Jupyter notebook.

Note: Considering the project report could be long, please generate your report in a pdf file and upload it on Canvas with the Jupyter notebook code.