Financial Time Series Prediction using Recurrent Neural Networks

Deep Learning Project Proposal

Dhruv Batheja, 4716523

Tiamur Khan, 4247329

Kanav Anand, 4712870

Problem

The problem that we will be investigating is that of stock price prediction. This problem is interesting to us due to the real-world applications of stock price prediction as well as the challenge that is involved with dealing with complex, volatile and dynamic stock price time series. We will investigate the additional value of long short-term memory units (LSMTs) to recurrent neural networks (RNNs) for stock price prediction. If time permits, we will also investigate the performance of the model trained on stock prices to other time series data sets such as weather, house prices, energy prices etc.

Data

We plan on using the New York Stock Exchange data-set from Kaggle. The prices in the data-set were obtained from Yahoo Finance, the fundamentals from Nasdaq Financials and it was further extended by some fields from EDGAR SEC databases.

Method

Due to the sequential nature of the problem, recurrent neural networks (RNNs) with long short-term memory units (LSTMs) will be used to solve the problem. LSTMs are an enhancement to RNNs, and enable them to memorize long range dependencies. We plan to implement pretrained models using Keras and TensorFlow.

Reading

[1], [2], [3], [4]

Research question

1. Can we predict stock price movements using recurrent neural networks?
2. What is the additional value of LSTM to RNNs for stock price prediction?
3. Can we extend these model for other time series data sets such as weather forecasting, house price predictions, energy price forecasting etc?

Evaluation

1. We will split the data-set into training and test data. The ground-truth prices in the test set and the prices predicted by our neural network will be presented in a single plot. We expect that our neural network is able to follow the trends of the ground-truth data. The mean squared error will be used to measure performance.
2. We will investigate the precision and time complexity of RNNs with LSTM and compare it with RNNs. Again, mean squared error will be used to measure performance.

3. We will run our model on data sets from different domains and evaluate its performance using mean-squared error metric.

Timeline

We plan to meet at least twice a week and work on the project together. Timelines for the major tasks can be found below:

Week	**Task**
Week 1	Reading the Literature
Week 2	Basic setup & Preprocessing the features
Week 3	Implementing a simple working RNN model
Week 4	Working on Milestone report and backlogs
Week 5	Implementing LSTM
Week 6	Fine tuning and optimizing hyperparameters
Week 7	Trying out the model on other timeseries datasets & Poster
Week 8	Final report, Poster and backlogs

Feasibility

To maintain feasibility of the project we will initially focus on simple recurrent neural networks only. Thereafter, we plan on implementing LSTMs as a performance enhancing technique. If there is additional time we may consider implementing other enhancement techniques. We will also demonstrate our model on a different time series dataset without any further tuning to have an overview of the robustness of the model.

References

[1] Samuel Edet. Recurrent neural networks in forecasting s&p 500 index. 2017.

[2] C. Krauss T. Fischer. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 2018.

[3] Rao Y Bao W, Yue J. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE*, 2017.

[4] Jurgen Schmidhuber Sepp Hochreiter. Long short-term memory neural computation. *MIT Press*, 1997.