Multi-Layer Perceptron (MLP), and Long-Short Term Memory (LSTMs) for stock price forecasting

Cheng-Lin Li
University of Southern California
April 26, 2017

Abstract

The one of the most challenging issue is stock price or indices prediction in the financial industry. On the other hand, machine learning and big data techniques in vision recognition has matured considerably over the last decade. This research adopts Multi-Layer Perceptron (MLP) and Long-Short Term Memory (LSTMs) neural networks to compete with Dynamic-radius Species-conserving Genetic Algorithm (DSGA) for short term stock price prediction. The result indicates that MLP may have a better potential than DSGA on short term stock price prediction and that LSTMs may require more training data to surpass DSGA.

Keywords - Stock Market, Machine Learning, Neural Networks (NN) for finance, Time Series, Multi-Layer Perceptron (MLP), Long-Short Term Memory (LSTMs), Prediction

1. Introduction

The stock price prediction is the ultimate goal in the financial industry. A good investment decision includes right targets with the correct timing. However, it is so difficult to predict the stock price accurately in reality because all factors which impact a stock price are both non-stationary and multiple dimension time series data [1]. Since the stock price is arbitrary, predicting stock price is a challenging task and few experts and economists can completely understand how the price will be. Two major approaches were adopted, fundamental analysis and technical analysis. Fundamental analysis contrastively recognizes the pattern of a stock price trend by focusing on the value of stocks, the potential of industries, economic situations, and political climates to suggest the future. Technical analysis evaluates the pattern of stock past prices and trading volumes to suggest the future prices [1].

In the past decade, technical analysis experienced a dramatic growth due to the maturity of big data processing and growth of computing power. Genetic Algorithms (GA), Artificial Neural Networks (ANN), Hidden Markov Models (HMM), and Support Vector Machine (SVM) are introduced for technical analysis of stock price prediction by long term data calculation [2]. However, the historical trend of the stock price may be less similar to the circumstances of today. The goal is based on less than one year Dow Jones stock information to predict which stock will reach the highest price increasing rate in the coming weeks. This study not only combines short term stock price information and the data of the economic environment to improve the prediction results, but also adopts the latest research results on vision recognition [3] to apply to the financial domain. This study introduces Multi-Layer Perceptron (MLP), and Long-Short Term Memory (LSTMs) to deal with the challenging goal. A Dynamic-radius Species-conserving Genetic Algorithm (DSGA) [4] was chosen as a baseline for comparison.

2. Potential Solutions

The significant improvements in vision recognition introduce neural network algorithm as a candidate in feature extraction in image or video [3]. It proves the capacity of neural network, especially LSTMs. On the other hand, another paper leverages weekly based market closing prices for 6 months and conducts Dynamic-radius Species-conserving Genetic Algorithm (DSGA) for financial forecasting of Dow Jones Index Stocks [4]. In this research, the trend of stock market indices is treated as a “diagram”, and DSGA is chosen to be the baseline for comparison. The main aim of this research is leveraging the power of multilayer artificial neural network to predict the trend of a stock price. Based on similar data set, a neural network was trained to learn the pattern from the price and dividend data of Dow Jones stocks from historical data to predict which stock will provide the maximum price percentage change in coming weeks. This research adopted two individual neural networks, MLP and LSTM to perform the experiments and compare the results with the baseline algorithm.

2.1 Multilayer Perceptron (MLP)

A feedforward and backpropagation linear calculation mathematic model based on artificial neural network model (ANN) was established by W. S. McCulloch and W. Pitts [2]. MLP model maps sets of data between inputs and outputs by the multilayer hidden perceptron in figure 1. Each neuron in multilayer perceptron usually feeds with weighted inputs, and applies linear or nonlinear activation function to generate the output [5]. A two hidden layers MLP are constructed for this project.

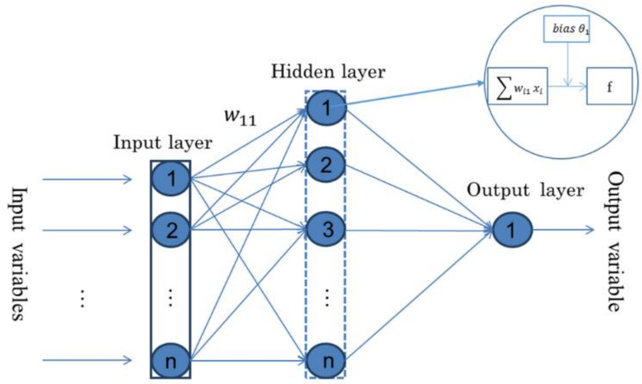

Figure 1. The architecture of the Multilayer Perceptron network [5].

2.2 LSTM

Long short-term memory algorithm is an enhanced architecture of the recurrent neural network (RNN) model. An LSTM network is programmed by the proper weight matrix and incorporates memory units to learn from experience to perform classification, processing and time series prediction tasks [3].

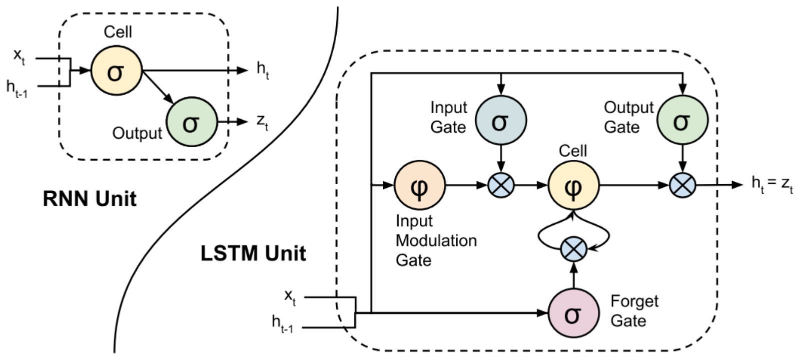

Figure 2. The architecture of RNN & LSTM [3].

This research constructs an LSTM with 256 cells in the hidden layer, and a basic LSTM Cell with "forget bias" is set to 1.0, and a linear activation is used for a RNN inner loop last output.

2.3 Dow Jones Index

Dow Jones Index is commonly known as a stock market index that created by Charles Dow, with the index first (non-industrial) published on February 16, 1885. The scaled average index is price-weighted and compensates for the effects of stock splits and other adjustments to generate a consistent value.

3. Activity

The weekly stock data for Dow Jones Index comprises data reported by the major stock exchanges. This dataset was first used in Brown, M. S., Pelosi, M. & Dirska, H. (2013) [4] for the Financial Forecasting of Dow Jones Index Stocks. In our research, each record not only contains stock prices during the week, but also includes the percentage of return that the stock has in the following week as the labeled data for training. This can help us to train and test the capability of prediction in our models. The quarter 1 (Jan-Mar, 2013) data was adopted for training and quarter 2 (Apr-Jun) data was adopted for testing purpose. Aside from the original data set, two additional data sources which include ISM manufacturing index and Personal Consumption Expenditure (PCE) price index, were introduced to represent the circumstance of external environments.

4. Results

4.1 MLP

a. The Network Structure of MLP as below:

Cost function: sigmoid cross entropy with logistics
Optimizer: Gradient descent
Activation function: tanh.

b. Parameter Settings

Table 1 lists parameters that are configured on the model. All parameters are determined during numbers of experiments.

Table 1. Parameters settings in MLP

Parameter	Definition	Value
learning_rate	Learning rate	0.01
training_epochs	Iteration times	1000
n_random_seed	Seed to generate the random value list	4
n_mean	Mean of initial Weights & bias term	0
n_stddev	Standard deviation of initial Weights & bias term	1
n_hidden_1	Number of perceptron in first hidden layer	400
n_hidden_2	Number of perceptron in second hidden layer	200

c. Selected Stocks

After the training, the model predicted the best stock for next 13 weeks as table 2 displays.

Table 2. Stocks selected in 13 weeks by MLP

Week	Stock	Return	Week	Stock	Return
1	BA	3.76%	8	CVX	1.57%
2	AXP	-0.24%	9	CVX	-5.43%
3	KO	0.91%	10	CVX	-1.03%
4	CVX	1.98%	11	CVX	-0.92%
5	AXP	1.68%	12	DD	2.30%
6	AXP	-1.14	13	AXP	7.94%
7	AXP	3.58%			

d. Return rates

To indicate the effectiveness of this method, several other well-known methods or indicators are compared. DSGA is the method from the paper referred to a baseline. Dow Jones Index shows the rate of return by investing an equal amount of money in each of the 30 stocks [4]. Maximum is the actual optimal return rates.

Table 3. MLP results compared against other methods [4]

Methods	Quarter	Week
MLP	14.96%	1.25%
DSGA	7.075%	0.54%
Maximum	51.82%	4.32%
Dow Jones Index	1.66%	0.13%

e. Discussion of Results

Comparing the data of table 3, MLP method beats the DSGA and Dow Jones Index in a definite advantage. MLP provides a good weekly return rates as 1.25%, almost 10 times to Dow Jones Index's performance.

However, the result is not actually reliable. The control factor in this result may be the n_random_seed, which is set to be 4 after several trials. With this value, the accuracy of prediction seems to keep at around 23.08% even without a well-trained neural network. Since the MLP is trained by very limited data set, it is difficult to well-trained the network. More training data may stabilize the MLP model to predict the real world.

f. Overfitting

The measures of overfitting display in figure 3. The axis of the diagram is Epoch and the y-axis is the training costs. As data presents, owing to our implementation, the training cost is obviously declining with the increment of Epoch. The training that the epoch overs 2000 times is treated as overfitting.

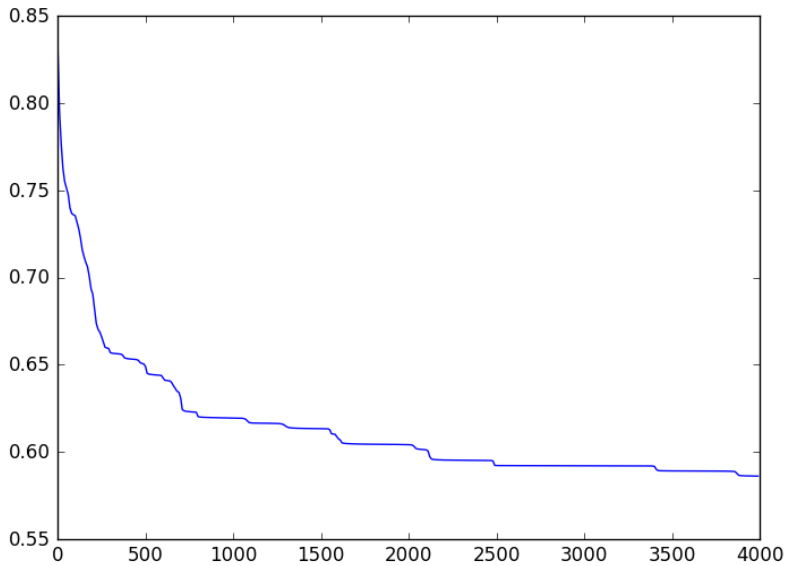

Figure 3. The costs versus epochs diagram

4.2 LSTM

a. Network Structure:

Cost function: softmax cross entropy with logistics

Optimizer: Adam algorithm

Activation function: tanh

b. Parameters

Table 4 is parameters that are configured on the LSTM model. These parameters are determined during numbers of experiments.

Table 4. Parameters settings in LSTM

Parameter	Definition	Value
learning_rate	Learning rate	0.001
training_iters	Iteration times	3000

n_random_seed	Seed to generate the random value list	4
n_mean	Mean of initial Weights & bias term	0
n_stddev	Standard deviation of initial Weights & bias term	1
n_input	Data input of record number	15
n_steps	Time steps	30
n_hidden	Number of perceptron in first hidden layer	256
n_classes	Classes (0-29 stocks)	30

c. Selected Stocks

Table 5. Stocks selected in 13 weeks by LSTM

Week	Stock	Return	Week	Stock	Return
1	JPM	-0.35%	8	JPM	0.11%
2	CVX	-1.01%	9	JPM	-1.94%
3	JPM	0.03%	10	JPM	-0.30%
4	DIS	1.96%	11	JPM	2.60%
5	MMM	-0.69%	12	JPM	-0.58%
6	CVX	-3.69	13	JPM	5.08%
7	JPM	2.35%			

d. Rates of return

Table 6. LSTM results compared against other methods [4]

Methods	Quarter	Week
LSTM	3.55%	0.30%
DSGA	7.075%	0.54%
Maximum	51.82%	4.32%

Dow Jones Index	1.66%	0.13%

e. Results

The result of return rate is 3.55% a quarter which is better than Dow Jones index performance (1.66%) but worse than DSGA algorithm (7.075%) that was chosen as the baseline.

f. Overfitting

Due to the cost / loss function cannot present the correct results, the measures of overfitting cannot be performed for LSTM. This task requires more resources to determine the results.

5. Conclusion and Future work

In this research, both MLP and LSTM cannot suggest an outstanding performance compared with vision recognition or speech prediction. It may be caused by lack of training data and improper initial weighting assignment on the network.

Data is king. Neural network especially need abundant data during the training period to generate a relatively reliable weighted network to perform the trend prediction of the stock price. LSTM especially requires more memorized data to generate better model. In the meantime, a proper initial weighting term leads to a better result in this research, but the final weighted network needs more testing in the real-world environment to prove the accuracy of the prediction. Parameters adjustment is an art in neural networks. It requires plentiful resources which include data and time to train a neural network on specific domains for precise results.

Aside from the technical analysis data, the MLP and LSTM may be benefitted by incorporation of fundamental analysis data into the model. Not only stock price but also macro-economic data may lead the neural network to a more comprehensive real world model to predict the trend in the future. On the other hand, it is still unknown to combine the power of MLP and LSTM on financial forecasting, and the combination of two neural networks may improve the performance of the stock price prediction. Above, all suggestions require more study to be clarified in the future.

References

1. J. Patel et al. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst. Appl. 42(1), pp. 259-268. 2015.
2. W. Ahmad. Analyzing different machine learning techniques for stock market prediction. International Journal of Computer Science and Information Security 12(12), pp. 17-23. 2014. Available: http://search.proquest.com.libproxy1.usc.edu/docview/1665221106?accountid=14749.
3. J. Donahue et al. Long-term recurrent convolutional networks for visual recognition and description. Presented at 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015, Available: http://dx.doi.org/10.1109/CVPR.2015.7298878. DOI: 10.1109/CVPR.2015.7298878.
4. M. S. Brown, M. J. Pelosi and H. Dirska. Dynamic-radius species-conserving genetic algorithm for the financial forecasting of dow jones index stocks. Presented at 9th International Conference, MLDM 2013. 2013, Available: http://dx.doi.org/10.1007/978-3-642-39712-7_3. DOI: 10.1007/978-3-642-39712-7_3.
5. M. Qiu and Y. Song. Predicting the direction of stock market index movement using an optimized artificial neural network model. PLoS One 11(5), pp. n/a. 2016. Available: http://search.proquest.com.libproxy1.usc.edu/docview/1790065467?accountid=14749. DOI: http://dx.doi.org.libproxy1.usc.edu/10.1371/journal.pone.0155133 .

Bio-Statement

Cheng-Lin Li is a graduate student pursuing a Master's Degree in Data Informatics in the Viterbi School of Engineering at the University of Southern California. He received his Bachelor's Degree in Information Management at National Taiwan University. He was the Vice President of J.P. Morgan Asset Management (Taiwan) leading e-Commerce and Fund Administration Team in IT Application Development Department for 16.5 years on international projects. Prior to that, he worked for 3.5 years at Acer Inc. Taiwan as a Project Engineer leading internet and intranet websites build-up in the Taipei Application Center. His professional interests include artificial intelligence, especially focus on machine learning in financial areas, application development, and system securities.