Machine Learning in Finance: Why You Should Not Use LSTM's to Predict the Stock Market

Andrea Leccese

January 2019

Executive Summary

Recently there has been much development and interest in machine learning, with the most promising results in speech and image recognition. This research paper analyzes the performance of a deep learning method, long short-term memory neural networks (LSTM's), applied to the US stock market as represented by the S&P 500. The paper shows that, while this technique may have had good success in other fields like speech recognition, it does not perform as well when applied to financial data. They are in fact characterized by high noise-to-signal ratio, which makes it difficult for a machine learning model to find patterns and predict future prices.

The paper is structured as follows. The first section introduces LSTM's and why they may be applied to financial time series. The second section analyzes the performance of an LSTM applied to the S&P 500. The third section concludes.

Contents

About the Author

Andrea is President and Portfolio Manager at Bluesky Capital. Prior to founding Bluesky Capital, Andrea worked as a Quant at Sauma Capital, a quantitative hedge fund in New York, and as a management consultant at Ernst & Young. Andrea holds an M.S. in Financial Engineering from Columbia University, and an M.S. in Industrial Engineering from the University of Rome "La Sapienza".

Contact

For more info about our investment products or questions about this paper, you can use the contact information below:

T: +1 (646) 465-9295
E: info@blueskycapitalmanagement.com
W: www.blueskycapitalmanagement.com

1. What is an LSTM?

Long short-term memory (LSTM) neural networks are a particular type of deep learning model. In particular, it is a type of recurrent neural network that can learn long-term dependencies in data, and so is usually used for time-series predictions.
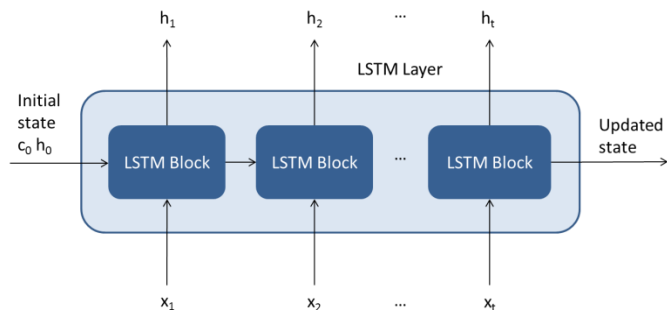
Figure 1 shows the architecture of an LSTM layer.

Figure 1 Architecture LSTM layer

As the figure shows, it is composed of a repeating core module. This structure makes the LSTM capable of learning long-term dependencies. The first LSTM block takes the initial state of the network and the first time step of the sequence X_1, and computes the first output h1 and the updated cell state c_1. At time step t, the block takes the current state of the network (c_{t-1}, h_{t-1}) and the next time step of the sequence X_t, and computes the output h_t and the updated cell state c_t.
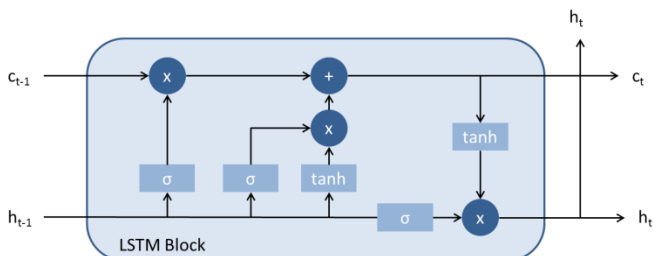

Figure 2 Architecture LSTM block

Figure 2 shows the architecture of an LSTM block, which is composed of 4 main components:

- **Input gate**: it controls the level of cell state update

- **Forget gate**: it controls the level of cell state reset

- **Cell candidate**: add information to the cell state

- **Output gate**: it controls the level of cell state added to the hidden gate.

In order to learn the features of the modeled task and be able to predict, an LSTM needs to be trained. This process consists in computing the weights and biases of the LSTM by minimizing an objective function, typically RMSE, through some optimization algorithms. Once the model it's trained on an initial training dataset and validated on a validation set, it is then tested on a real out of sample testing. This ensures that the model did in fact learn useful features and it is not overfitted on the training set, with poor prediction capabilities on new data. The next section analyses the performance of an LSTM applied to the S&P 500.

2. Performance LSTM applied to the US equity market

The used dataset is composed of closing daily prices for the US stock market, as represented by the S&P 500, from January 3, 1950 to January 4, 2019, for a total number of 17,364 observations. The data is divided in 60% for training, 20% for validation, and 20% for testing.

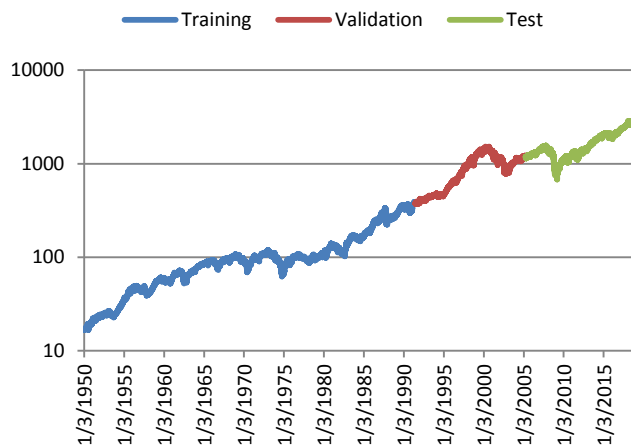Figure 3 shows the data used for the analysis on a log scale.

Figure 3 S&P 500 close prices (log scale)

In our analysis we trained an LSTM neural network composed of 1 hidden layer, 20 neurons, and time series length of 20 values. We tried different combinations for the neural network parameters and architectures and found similar results.

Figure 4 shows the actual prices compared to the next day values predicted by the trained LSTM on a log scale.
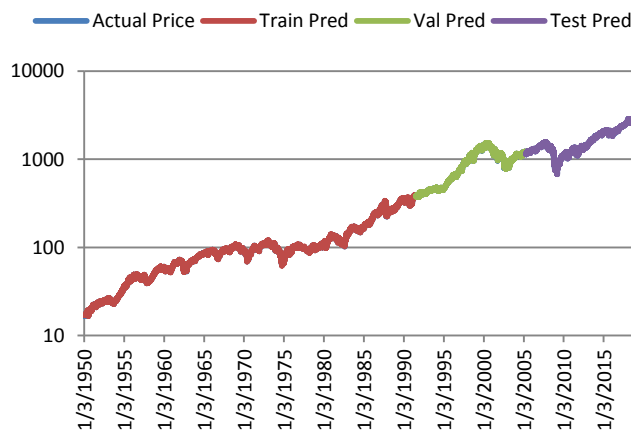

Figure 4 LSTM predicted vs actual price for the S&P 500 (log scale)

As it can be seen from the figure, the predicted values are very close to the actual prices, so that the underlying actual price cannot be seen properly. Thus, it may seem that the LSTM does a good job at predicting the next value for the time series under consideration.

Table 1 reports the performance stats for the LSTM accuracy in predicting the next day price.

Dataset	RMSE	R^2
Training	0.6227	0.8293
Validation	0.7326	0.7508
Test	0.8211	0.6976

**Table 1 LSTM prediction accuracy stats**

This data corroborates what we can see from Figure 4. The low values in RMSE and decent values in R^2 show that the LSTM may be good at predicting the next values for the time series in consideration.

Figure 5 shows a sample of 100 actual prices compared to predicted ones, from August 13, 2018 to January 4, 2019.
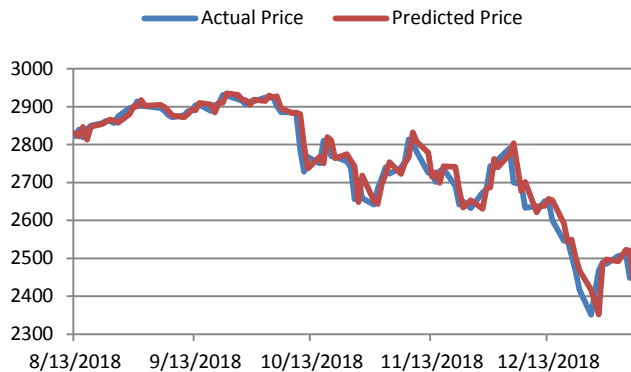
Figure 5 LSTM predicted vs S&P 500 price for 100 days

This figure makes us draw a different conclusion. While in aggregate it seemed that the LSTM is effective at predicting the next day values, in reality the prediction made for the next day is very close to the actual value of the previous day. This can be further seen by Figure 6, which shows the actual prices lagged by 1 day compared to the predicted price.
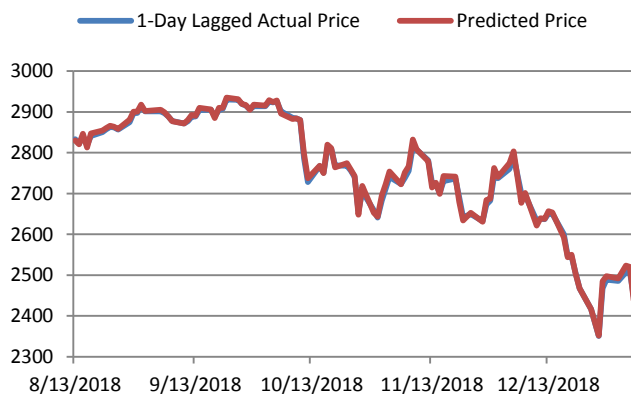
Figure 6 LSTM predicted vs 1-day lagged S&P 500 price for 100 days

As the figure shows, the 2 series are almost identical, confirming our previous conclusions.

These results demonstrate that LSTM is not able to predict the value for the next day in the stock market. In fact, the best guess the model can make is a value almost identical to the current day's price.

3. Conclusion

While it is true that new machine learning algorithms, in particular deep learning, have been quite successful in different areas, they are not able to predict the US equity market. As demonstrated by the previous analyses, LSTM just use a value very close to the previous day closing price as prediction for the next day value. This is what would be expected by a model that has no predictive ability.

This also highlights that while some machine learning techniques may be useful in finance, quantitative hedge funds must take another route and come up with strategies capable of delivering alpha for their clients.

Disclaimer

This report is prepared and circulated for informational and educational purposes only and is not an offer to sell or the solicitation of an offer to buy any securities or other instruments. The information contained herein is not intended to provide, and should not be relied upon for investment, accounting, legal or tax advice. This document does not purport to advise you personally concerning the nature, potential, value or suitability of any particular sector, geographic region, security, portfolio of securities, transaction, investment strategy or other matter. No consideration has been given to the specific investment needs or risk-tolerances of any recipient. The recipient is reminded that an investment in any security is subject to a number of risks including the risk of a total loss of capital, and that discussion herein does not contain a list or description of relevant risk factors. As always, past performance is no guarantee of future results. The recipient hereof should make an independent investigation of the information described herein, including consulting its own tax, legal, accounting and other advisors about the matters discussed herein. This report does not constitute any form of invitation or inducement by Bluesky Capital to engage in investment activity.

The information contained herein reflects, as of the date hereof, the views of Bluesky Capital Advisors, LLC (together with its affiliates, "Bluesky Capital") and sources believed by Bluesky Capital to be reliable. No representation or warranty is made concerning the accuracy of any data compiled herein. In addition, there can be no guarantee that any projection, forecast, or opinion in these materials will be realized. Certain information contained in the presentation discusses general market activity, industry or sector trends, or other broad-based economic, market or political conditions and should not be construed as research. The views expressed herein may change at any time subsequent to the date of issue hereof. These materials are provided for informational purposes only, and under no circumstances may any information contained herein be construed as investment advice or an offer to sell or a solicitation of an offer to purchase (or any marketing in connection thereof) any interest in any investment vehicles managed by Bluesky Capital or its affiliates. The information contained herein does not take into account your particular investment objectives, financial situations, or needs, and you should, in considering this material, discuss your individual circumstances with professionals in those areas before making any decisions. Alternative investments can be highly illiquid, are speculative, and may not be suitable for all investors. Investing in alternative investments is only intended for experienced and sophisticated investors who are willing to bear the high economic risks associated with such an investment. Investors should carefully review and consider potential risks before investing. Bluesky Capital may make investment recommendations and decisions that are contrary to the views expressed herein, and may sponsor and hold interests in investment vehicles that have holdings that are inconsistent with the views expressed herein.

Bluesky Capital makes no representations, express or implied, regarding the accuracy or completeness of this information, and the recipient accepts all risks in relying on this report for any purpose whatsoever. This report is being furnished to the recipient on a confidential basis and is not intended for public use or distribution. By accepting this report, the recipient agrees to keep confidential the existence of this report and the information contained herein. The recipient should not disclose, reproduce, distribute or otherwise make available the existence of and/or all or any portion of the information contained herein to any other person (other than its employees, officers and advisors on a need-to-know basis, whom the recipient will cause to keep the information confidential) without Bluesky Capital's prior written consent. This report shall remain the property of Bluesky Capital and Bluesky Capital reserves the right to require the return of this report at any time.

This paper is intended for academic and educational purposes and is not an investment recommendation. The hypothetical models discussed in this paper do not reflect the investment performance of any actual product or strategy in existence during the periods tested and there is no guarantee that if such product or strategy existed it would have displayed similar performance characteristics. A decision to invest in any product or strategy should not be based on the information or conclusions contained herein.

Some of the images, logos or other material used herein may be protected by copyright and/or trademark. If so, such copyrights and/or trademarks are most likely owned by the entity that created the material and are used purely for identification and comment as fair use under international copyright and/or trademark laws. Use of such image, copyright or trademark does not imply any association with such organization (or endorsement of such organization) by Bluesky Capital, nor vice versa.

Bluesky Capital

T: +1 (646) 465-9295
E: info@blueskycapitalmanagement.com
W: www.blueskycapitalmanagement.com