Disproving the specious efficiency presented by BiLSTM models

Hussain Manasi[1] and Dr. Vijay Bhaskar Semwal[2]

[1,2] *Maulana Azad National Institute of Technology, Bhopal*

Abstract— There is a significant amount of misleading trust when it comes to reliability on neural networks when it comes to predicting time series data like stock data. Apart from surface level faults like missing out on integration of fundamental analysis, it also misunderstands many principal concepts like scaling, inefficient data, and in our case, overpowered models. This paper aims to tackle the particular misconception of efficiency when it comes to predictions made by BiLSTM models on stock data and why any results obtained by the BiLSTM model may in fact be prone to overfitting and hence altogether be wrong.

Keywords— LSTM, BiLSTM, Univariate, Multivariate, Time Series Analysis

I. Introduction

Quite simply put Time Series analysis refers to studying data collected in a chronological manner from which insights can be brought out to predict future trends. Time series data can be recorded in regular time intervals or non linearly for other more subtle applications.[1][2][3][4] Stock data is an application of Time Series Analysis that has the characteristic of being significant in a time bound manner. The problem statements of time series analysis have effectively been evaluated by the introduction of the LSTM, or Long Short Term Memory. This RNN architecture consisting of gates and activation functions within each cell have allowed for improvement over the initial errors when simple RNNs were implemented, mainly the problem of vanishing gradients, and on top of that providing a method to perform time series analysis while ensuring significant featuers of attributed are maintained or 'remembered'. The BiLSTM is a further technique built upon the LSTM which introduces the idea of reading input data not only unidirectionally but bidirectionally as well. What this means is that to predict a data point $y(t)$, the simplest implementation of LSTM would evaluate the data points of $y(t-1), y(t-2), y(t-3)$.... But a bidirectional LSTM takes the liberty to do as stated before, and then also process data points $y(t+1), y(t+2), y(t+3)$... if provided in the training set.[5]

Combining all of the above provided context, there is in the research space, a whole lot of content regarding predicting prices in the stock market and how best to go about it. This sort of research is still prevalent even now despite the clear understanding that models trained on certain training data will not provide reliable predictions, and further if the predictions are accurate, they are most probably contrived enough to be absolutely useless in the real world. Several factors initially haunt the feasibility of proper stock price predictions mainly:[6][7]

1. **Insufficient Datasets**: Insufficient datasets point to results obtained by tweaking the model to fit the situation and gives less generalised results. The requirement of not only a training and testing dataset but also validation and cross validation has proven to be crucial to proper recommendable models.

2. **Inappropriate Scaling and Measures**: Scaling the input data is a common practice that helps bring down the range of values that the model has to learn on. However, performance can tend to be poor in scenarios where scaling leaves out data that wasn't initially visible in the training dataset. If the range of the testing set exceeds that of the training set then the predictions may hit a peak possible value. This can of course be handled by making sure that a sufficient range is seen by the model, but that makes the model vulnerable to more error. Additionally, the measures used in quantifying the performance of the model like the RMSE or MAE, can prove to be quite misleading. Not only do they not reflect a good understanding of the trend that a stock is following (or going to follow for that matter), it can also lead to false interpretation of a model's performance from a researcher.

3. **Time Series Tracking**: Some model performances can seem too good to be true, presenting error rates so likeable, no one could resist using it for personal gains. However, the results mimic the time series data itself, and that especially becomes the case when model aims to predict next day's price rather than sequence to sequence prediction model. This can be looked after by first considering elimination of using prices as inputs and outputs altogether. A stock market broker probably focuses on returns rather than exact prices anyway, and modern financial markets do also focus on strategy of decided whether to hold, sell, or buy for this exact reason. Price Returns is a better metric to consider when working with stock predicting models.

Now, to further dive into the models we are going to use. Without delving too deep into the architecture of an LSTM cell, the gist is that an LSTM can remember important trends similar to how we humans remember certain moments in our lives better than others. The RNN concept is to train the model by repeatedly pushing the same data through the cells and varying the weights to find the most optimal solution. LSTMs have since their inception been quite appropriately applied to the stock market for acquiring reliable predictions and is a good starting point and name brand when it comes to application of RNNs to stock price predictions.[8][9][10][11][12] The BiLSTM was built as a modification to improve context based predictions. A perfect implementation of the BiLSTM is observed in text completion. A certain sequence of text can be predicted by the model given it is trained on similar complete sequences. However, BiLSTM have been used by certain researchers to predict stock prices. As stated above, the practice of predicted stock prices is flawed in and of itself anyway, but additionally the BiLSTM training considers not just the past prices of a stock but also the future prices? This sounds like an absurd method to train a model for a time series application like stock data that has clearly been built depending on only its past values and where the future values haven't even been conceived at that point. [13][14][15][16]

Hence, this paper especially focuses on the misuse of BiLSTM on stock data. The experiment is especially elaborated in section II but as a summary; we have trained a univariate BiLSTM model and a multivariate stacked LSTM to find that that the univariate BiLSTM is subject to not only overfitting but also unrealistic results that should not be possible given some of the faults described earlier.

II. Experiment Setup

The dataset used is that of ADANIPORTS from the NIFTY50 dataset and an 80-20 split is performed on the dataset. The publicly available data can be accessed from both Kaggle, or from the Yahoo Finance API. The earliest date of the dataset goes back to 2007-11-27, and the last date is that of 2021-04-30. It is common knowledge that such a long time span would not be useful and would only confuse the model further by incorporating into consideration the 2008 great depression as a factor, and the intermediate years which would provide little impact on the stock prices we may see today. Therefore, any data points before 2014-01-02 are cut out from the entire experiment.

The model summaries are as follows:

1. **The Univariate BiLSTM**: Consisting of an initial layer of 20 bidirectional neurons followed by 10 more bidirectional neurons and finally ending with a single dense layer;

```
Total params: 7,621
Trainable params: 7,621
Non-trainable params: 0
```

2. **Multivariate Stacked LSTM**: This model iteration made significant use of the dropout layers sandwiched between 4 layers of LSTM cells each having 100,100,100, and 50 neurons each respectively. The parameters of the model are as follows;

```
        Total params: 233,451
Trainable params: 233,451
Non-trainable params: 0
```

III. Results

Shown below are the resulting graphs obtained using our prediction models. Fig. 1 is a full depiction of the whole dataset. Fig. 2 is predictions made on the testing set by Univariate BiLSTM, and Fig.3 is predictions made on the testing set by Multivariate Stacked LSTM.

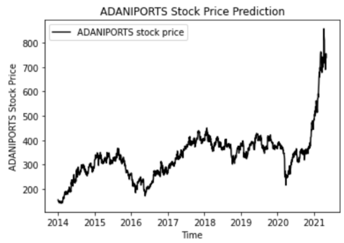

Fig. 1: The full dataset being used in both the models

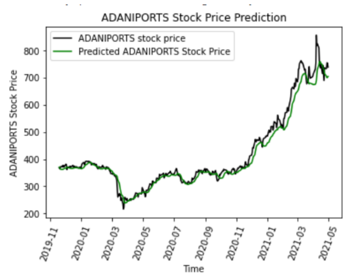

Fig. 2: Predictions made by Univariate BiLSTM

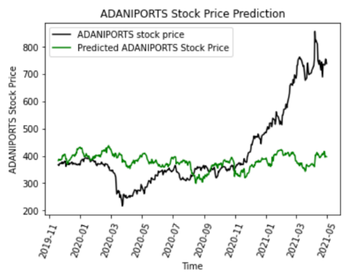

Fig. 3: Predictions made by Multivariate Stacked LSTM

	MAE	MSE
Univariate BiLSTM	17.0	25.9
Multivariate Stacked LSTM	94.3	139.5

Table 1: Errors recorded for both experiments

IV. Discussions

As we can see, in fig. 2 the univariate BiLSTM model is performing at an alarmingly good measure. This is in reference to when we discussed time series tracking earlier. The model seeks to predict not next day price values based on the price value of 60 days before it, and ends up mimicking the trend presented in the testing set with an obviously faulty prediction, one that would perform well here, but would fail tremendously in real life scenarios. The multivariate stacked LSTM in contrast considers a larger piece of the picture and hence while it looks back 30 days before it, it refrains from exceeding logical limits when it comes to making sequence to sequence predictions. It is a much more realistic albeit high error result of the model performance.

The range of predictions made by both the experiments goes up to predicting 2 years worth of stock price from the end of the training set. While at first this may seem ridiculous as anything can happen that would drastically influence the stock market in a span of 2 years, however, that is the intent. The idea of predicting stock prices should span to more than just how high or low the price will be tomorrow or the day after, or in the next week. Sure, anything can happen in 2 years time, but predicting to that far in the future can allow a skilled researcher to learn whether the model will perform positively or not in the long run. The dataset used of ADANIPORTS finds it difficult to reflect an upwards trend as the training data itself has constant price range of about where the predictions are being made. Using a dataset like TSLA would provide an upward trend in predictions but would still be limited due to inappropriate scaling, and reaching of out of bounds values.

V. Conclusions

This paper attempts to study the exact steps when it comes to stock prediction and scrutinize the possibility of misuse and unforeseen errors that creep into the experiments conducted by the researcher. One of which is the use of BiLSTM for stock prices. The learning process is viewed as entirely wrong, and the results are specious. It may be due to mistaken application, or by misguided need to get results but the practice of predicting the future by learning from the future of the past is presented as a foolish alternative if the aim is only to gain better results.

References

[1] "Time series - wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Time_series

[2] "What is time series data? | definition, examples, types uses." [Online]. Available: https://www.influxdata.com/what-is-time-series-data/

[3] "What is the difference between linear and non-linear time series? - cross validated." [Online]. Available: https://stats.stackexchange.com/questions/351950/what-is-the-difference-between-linear-and-non-linear-time-series

[4] "What is time series data? | definition, examples, types uses." [Online]. Available: https://www.influxdata.com/what-is-time-series-data/

[5] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2673–2681, 1997.

[6] "(pdf) common mistakes when applying computational intelligence and machine learning to stock market modelling." [Online]. Available: https://www.researchgate.net/publication/230715644_Common_Mistakes_when_Applying_Computational_Intelligence_and_MachineLearning_to_Stock_Market_modelling

[7] "Mistakes in stock prediction: Trying to predict the price." [Online]. Available: https://lazyprogrammer.me/mistakes-in-stock-prediction-trying-to-predict-the-price/

[8] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "Stock price prediction using lstm, rnn and cnn-sliding window model," *2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017*, vol. 2017-January, pp. 1643–1647, 11 2017.

[9] Y. Guo, "Stock price prediction based on lstm neural network: The effectiveness of news sentiment analysis," *Proceedings - 2020 2nd International Conference on Economic Management and Model Engineering, ICEMME 2020*, pp. 1018–1024, 11 2020.

[10] S. Goswami and S. Yadav, "Stock market prediction using deep learning lstm model," *2021 International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON 2021*, 2021.

[11] P. Srivastava and P. K. Mishra, "Stock market prediction using rnn lstm," *2021 2nd Global Conference for Advancement in Technology, GCAT 2021*, 10 2021.

[12] J. Kavinnilaa, E. Hemalatha, M. S. Jacob, and R. Dhanalakshmi, "Stock price prediction based on lstm deep learning model," *2021 International Conference on System, Computation, Automation and Networking, ICSCAN 2021*, 7 2021.

[13] S. Mootha, S. Sridhar, R. Seetharaman, and S. Chitrakala, "Stock price prediction using bi-directional lstm based sequence to sequence modeling and multitask learning," *2020 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2020*, pp. 0078–0086, 10 2020.

[14] J. Y. Lim, K. M. Lim, and C. P. Lee, "Stacked bidirectional long short-term memory for stock market analysis," *3rd IEEE International Conference on Artificial Intelligence in Engineering and Technology, IICAIET 2021*, 9 2021.

[15] J. Shah, R. Jain, V. Jolly, and A. Godbole, "Stock market prediction using bi-directional lstm," *Proceedings - International Conference on Communication, Information and Computing Technology, ICCICT 2021*, 2021.

[16] K. A. Althelaya, E. S. M. El-Alfy, and S. Mohammed, "Evaluation of bidirectional lstm for short and long-term stock market prediction," *2018 9th International Conference on Information and Communication Systems, ICICS 2018*, vol. 2018-January, pp. 151–156, 5 2018.