Empirical Evaluations of Time Series Forecasting Methods Applied On Stock Prediction

Ronghan Che
Punita Verma
Erwin Feuzeu
cherh63@vt.edu
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

ABSTRACT

There exists various approaches to conduct stock prediction in the current academic field. Those different approaches could be distinguished according to the training dataset, feature selection and extraction methods, neural network models, types of the technical encoders formed the pipeline, and the error evaluation techniques. On the other hand, time series forecasting models have been proved with good performance in both long term and short term stock price predictions. The main goal of our project is to conduct empirical evaluations between neural networks applying Long Short Term Model(LSTM) with different input methods.

ACM Reference Format:
Ronghan Che, Punita Verma, and Erwin Feuzeu. 2021. Empirical Evaluations of Time Series Forecasting Methods Applied On Stock Prediction. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

1.1 Problem Definition

Generally Long Short Term Models have decent prediction accuracy in time series forecasting. However, there exists factors which could influence the model's performances including the imported stock attributes, structure of the neural networks, feature selection approaches, and hyper-parameter tuning. We experimented with four data-sets: The first one is composed of the basic stock attributes including open price, close price, daily highest price, daily lowest price, and adjusted close price. The second dataset will incorporate more stock market indicators as extra attributes. The third dataset will use Principal Component Analysis(PCA) on the entire set of imported attributes to select out the principal components which takes the largest weight in the input matrix. The last dataset will reserve the basic five stock attributes(Open, Close, High, Low, Adj Close); and apply PCA on the rest of the technical indicators.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1.2 Motivation

Our previous approach showed that the implementation of the LSTM simply trained by daily close price even could achieve a decent accuracy measured by Mean Square Error(MSE) and Root Mean Square Error(RMSE). This gives us a motivation to enhance the model's performance by importing more stock attributes into the training data-set. On the other hand, there exists huge performance difference due to the models' hyper-parameters and lengths of the prediction periods. We want to find out the relations between those factors influencing the model's performance and their values.

2 RELATED WORK

2.1 Literature Survey

Stock prediction has been a popular topic in machine learning field; there exists numerous of approaches composed of different neural network structures incorporated with different input features(Weiwei 2020). This paper indicates that the combination of LSTM, technical indicators, and macroeconomics attributes has been a popular option according to its data availability and performance evaluation.
In (David, Nelson 2017), it introduced the structure of time series forecasting models, their related backgrounds, and the usage of Long Short Term Models to make stock predictions, alongside with technical indicators.
Alexander, Konstantinos, Michael et al.(2019) introduced Gluon Time Series (GluonTS), a library for deep learning based time series modeling. I introduced the necessary components and procedures to implement time series forecasting neural networks.
Konstantinos, Syamma et al.(2020) made a series of comparisions between different neural network architectures; and provided directions of hyper-parameter tuning in designing NN models.
In (Lan, Jorge 2020), it introduced a dimension reduction approach, principle component analysis including its background, math formulas, implementation approaches, and current applications.
(Tingwei, Yuting 2018) introduced a method to improve the performances of LSTM in stock prediction by importing technical indicators and applying dimension reduction techniques such as principal component analysis.

2.2 Limitations of Existing Approaches

Due to the nature of the stock market's volatility, the current biggest technical challenge for stock prediction is to tackle with the correlations between multiple influencing factors and analyze a general

pattern applied on a schema covering those factors. Currently, many regression models predicting the increase and decrease of a company's stock price can not achieve a prediction accuracy higher than 60 percents.
On the other hand, models' performances will be significantly diminished if the goal is to predict multiple stock attributes, to incorporate with multiple company's stock data, or to extend the prediction periods which supports longer term prediction.

3 PROPOSED APPROACH

3.1 Input variables

Stock indicators have been proved with good abilities to predict the movement of company's stock prices, detect market signals such as oversold and overbought, and support analyzing the market value of a company's stock dynamically according to multiple dimensions. The list of variables we imported in different input methods include the basic stock attributes including Open(OP), Close(CL), High(HI), Low(LO), Volume(VO), and Adjusted Close(AJC). Our approach to adapt the technical indicators is to first brutally include the entire list of the stock indicators(Table 1); then apply feature extraction techniques to keep only the indicators which have the biggest influence on the training data-set.

- Accumulation indicator(A/D) is a cumulative indicator that uses volume and price to assess whether a stock is being accumulated or distributed.
- Exponential Moving Average(EMA) is a technical chart indicator that tracks the price of an investment (like a stock or commodity) over time.
- Moving Average(MA) is a calculation used to analyze data points by creating a series of averages of different subsets of the full data set.
- Momentum is the speed or velocity of price changes in a stock, security, or tradable instrument. Momentum shows the rate of change in price movement over a period of time to help investors determine the strength of a trend.
- Rate of Change(ROC) is a momentum-based technical indicator that measures the percentage change in price between the current price and the price a certain number of periods ago.
- Weighted Moving Average(WMA) is a technical indicator that assigns a greater weighting to the most recent data points, and less weighting to data points in the distant past.
- Aroon is a technical indicator that is used to identify trend changes in the price of an asset, as well as the strength of that trend.
- Average True Range(ATR) is a volatility indicator that shows how much an asset moves, on average, during a given time frame.
- Bollinger Bands(BB) indicate the stock market's volatility.
- Comodity Channel Index(CCI) is a momentum-based oscillator used to help determine when an investment vehicle is reaching a condition of being overbought or oversold.
- Chande Momentum Oscillator(CMO) is a technical momentum indicator developed by Tushar Chande.
- Chaikin Oscillator(CO) is the difference between the 3-day and 10-day EMAs of the Accumulation Distribution Line.

Table 1: Stock Indicators, n = time period

Indicator	Formula
A/D	$k * a(t)$
EMA	$(V_t * \frac{s}{1+d}) + (V_t * \frac{s}{1+d})$
MA	$\frac{A_1+A_2+...A_n}{n}$
Momentum	$V - V_x$
ROC	$(c/v - 1) * 100$
WMA	$\frac{P_1*x*n+P_1*x*n+...P_n}{[n*(n+1)]/2}$
Aroon	$(n - l/n) * 100$
ATR	$(1/n) \sum_{i=1}^{n} TR_i$
BB	$MA(TP, n) - m * \sigma[TP, n]$
CCI	$\frac{TP-MA}{.015*MD}$
CMO	$\frac{sH-sL}{sH+sL} * 100$
CO	$(3 - dayEMAofADL)(10 - dayEMAofADL)$
DPO	$\frac{X}{2}$ +1 periods agoX period SMA
IC	$\frac{9-PH+9-PL}{2}$
IC	$\frac{26-PH+26-PL}{2}$
IC	$\frac{CL+BL}{2}$
IC	$\frac{52-PH+52-PL}{2}$
MFI	$100 - \frac{100}{1+MFR}$
MACD	$12 - PeriodEMA26 - PeriodEMA$
OBV	$OBV_{prev} + v$
Relative Strength Index	$RSI_{stepone} = 100 = [\frac{100}{(1+\frac{AG}{AL})}]$
UO	$[\frac{A_7*4+A_17*2+A_28}{4+2+1}] + 100$
VI	Last TVI + Volume
WPR	$\frac{High-CLO}{High-LO}$

- Williams Percent Range(WPR) is a type of momentum indicator that moves between 0 and -100 and measures overbought and oversold levels.

3.2 Feature Extraction - Principal Component Analysis

As we introduced before, the first step in our procedures of expanding the schema of stock attributes is to brutally add all the most widely used stock indicators into the training data-set. It will be inevitable to introduce noising data as well. The aim of principal component analysis is to help with removing the noisy attributes by computing the variances of each principle component to get rid of the attributes taking less weighted-correlations with the rest of the attributes.(Figure 1)

3.3 Input Methods

In the first turn of our experiment, we set the time window as 60 days. The training data-set we chose is the TSLA's stock prices from January 2014 to January 2020. The test data-set is composed of the same company's stock price from January 2020 to January 2021.
The first input method takes only the stock basic attributes(Open, Close, High, Low, Adj Close) without applying feature extraction or importing extra attributes.(Figure 2)
The second input method brutally added all the stock indicators we introduced. Some technical indicators such as Commodity Channel

```
array([5.04944384e-01, 2.14764640e-01, 5.89417592e-02, 4.12364677e-02,
       3.36429924e-02, 2.18433682e-02, 1.71950143e-02, 1.49192493e-02,
       1.33452825e-02, 1.12603009e-02, 1.00756334e-02, 9.12402198e-03,
       7.13297306e-03, 6.06599310e-03, 5.12104761e-03, 4.65865713e-03,
       4.22928575e-03, 3.58597678e-03, 3.31824296e-03, 2.81786895e-03,
       2.42289971e-03, 1.70680252e-03, 1.17371499e-03, 1.09871884e-03,
       9.60602192e-04, 9.13971860e-04, 6.78745215e-04, 5.45882853e-04,
       4.42134507e-04, 3.80393089e-04, 3.10098054e-04, 2.44255294e-04,
       1.96448251e-04, 1.81589305e-04, 1.76273825e-04, 1.59422605e-04,
       7.75098815e-05, 5.36805092e-05, 3.12108532e-05, 1.24842627e-05,
       9.81222001e-06, 1.85970947e-07, 1.58173620e-15, 1.58173620e-15,
       1.58173620e-15, 1.58173620e-15, 1.58173620e-15, 1.58173598e-15],
      dtype=float32)
```

Figure 1: Variance of Imported Principle Components

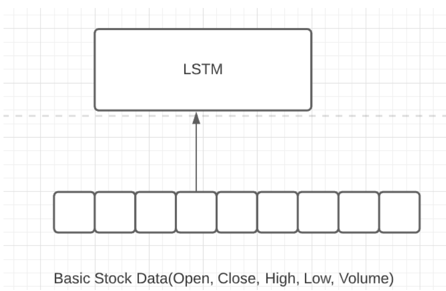

Figure 2: Input A

Index asks for a time period. Considering we are making predictions with a time window as 60 days, we added multiple CCI indicators each for 30, 40, 50 days. Eventually we imported 47 stock attributes in total including the basic stock data and the extra technical indicators.(Figure 3)

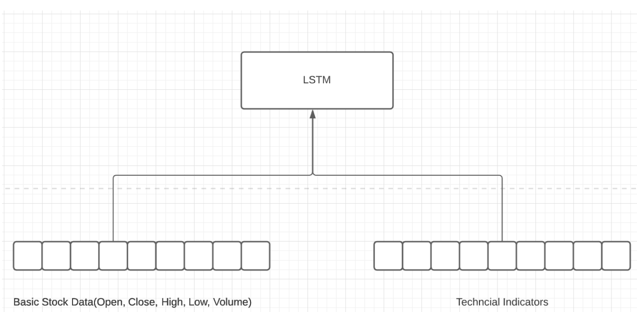

Figure 3: Input B

In the third input method, according to the result of the principal component analysis that applied on all the stock attributes including the five basic stock data and the extra technical indicators(Figure 1), we kept the first 24 principle components which are responsible of more than 90 percent variance.

In the last input method, we reserved the five basic stock attributes without applying PCA on them; then used PCA on the rest of the technical indicators. From the result of PCA that is a variance array orgnized in the same structure as (Figure 1) shows, we kept the 11 principle components taking up to 99 percent of the total variance. (Figure 5)

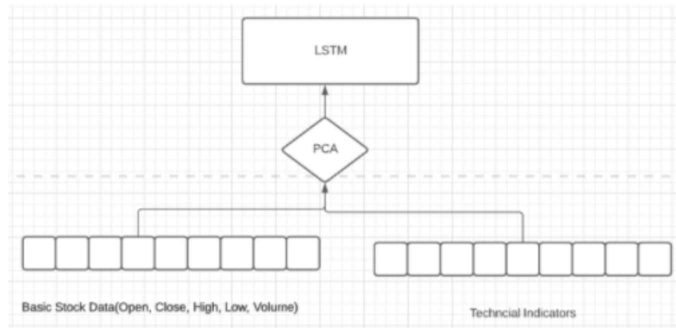

Figure 4: Input C

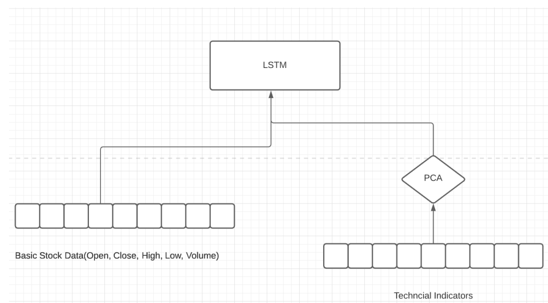

Figure 5: Input D

4 EXPERIMENTAL EVALUATION

4.1 Performance Comparisons of the Introduced Input Methods

Root Mean Square Error(RMSE) calculates the standard deviations of the prediction errors that can tell how "spread out" the data points are. Mean Absolute Error(MAE) can evaluate the difference between the prediction and the actual prices as well.

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(\frac{d_i - f_i}{\sigma_i}\right)^2}$$

$$mae = (\frac{1}{n})\sum_{i=1}^{n}|y_i - x_i|$$

We adopted those two methods to evaluate the performance of each model. The direct comparison between the actual price and prediction(Figure 6) and the error evaluations(Figure 7) are obtained. In a 60-day prediction period, surprisingly we observe that the first model having the simplest input data shows the best performance. Comparing between input method A and input method B, it can be observed that the prediction accuracy of input method B is significantly worse than input method A. Does it indicate that incorporating the extra stock indicators and applying the dimension reductions are effortless? The answer is absolutely not. We can also observe that input method D has improved the overall prediction accuracy than model C by applying more advanced input sequences.

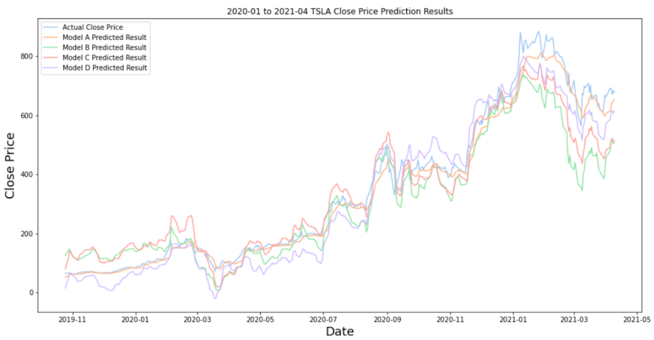

Figure 6: Prediction Result of Input A, B, C, D

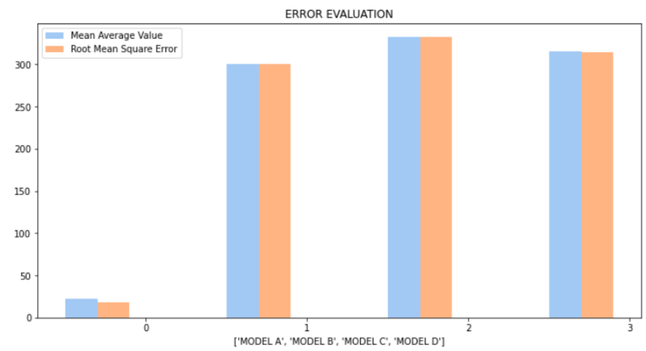

Figure 7: Error Evaluations of Input A, B, C, D

We think one possible reason caused the significant performance decrease is due to the technical indicators such as Bollinger Bands and Commodity Channel Index ask for specifying a time period, which we set for multiple values in the feature selection stage.Therefore, we made a few more experiments tested with different prediction periods. If you look at the bar chart(Figure 8),(Figure 9), the blue bars represent the mean absolute errors and the rmses of all models in case of making predictions in 50 days. The yellow bars represent the errors when predict stock price in 100 days. The green bars are the result in 200 days. And the red bars are in 300 days. It can be observed that as the length of the prediction periods increased, the performance will also get improved. So in a nutshell, long-term period looks having better performance than short-term prediction in our neural network model with time window as 60 days.

4.2 Initial Result

We have already discovered that the long-term predictions are doing better than short-term forecasting in this neural network. This is not normal; and we tried to find the related factors caused this result. In the time series forecasting problem, we have to predict a value at time T based on the data $T - N$ where N is the time window. My assumption for this question is that the length of the time window is related to the prediction accuracy of different lengths of prediction terms. My approach is to adjust the length of the time window to make predictions in a same length of time period. In the last set of experiment, I made error evaluations of setting time

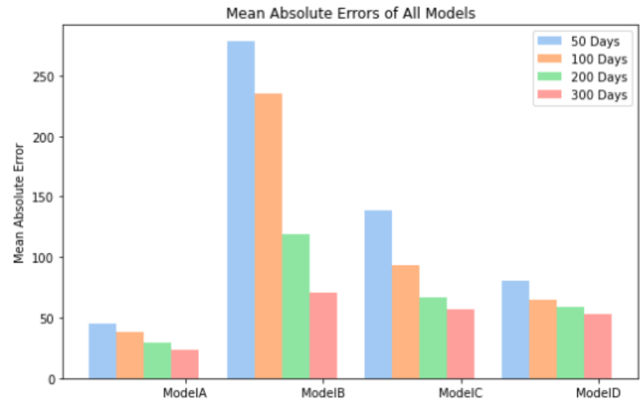

Figure 8: Mean Absolute Errors

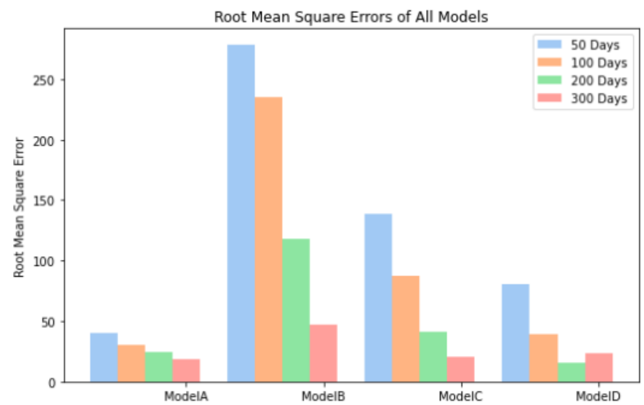

Figure 9: Mean Absolute Errors

windows equal to 60, 100, 150, 200 days, with the prediction period as 732 days to make Close price prediction. As the result(Figure 10) shows, the MAE and RMSE of the 100-day time window has the best performance. With that being said, time-window is also a factor we need to include in the list when we consider adjusting hyper-parameters to enhance the model's prediction accuracy.

Error Evaluation				
	Time Window	Prediction Period	MAE	RMSE
A	60	732	82.59	44.32
B	100	732	81.79	53.70
C	150	732	41.48	10.33
D	200	732	43.78	26.50

Figure 10

In a nutshell, factors influencing the same time series forecasting model include the imported attributes, dimension reduction methods, prediction terms, and time windows. You can check the full list of our experiment result including those factors. (Figure 11)

Error Evaluation				
	Time Window	Prediction Period	MAE	RMSE
A	60	50	45.01	39.78
B	60	50	277.96	277.96
C	60	50	138.71	138.72
D	60	50	80.55	80.55
A	60	100	37.92	30.39
B	60	100	235.37	235.37
C	60	100	93.30	87.71
D	60	100	65.17	39.19
A	60	200	29.23	24.18
B	60	200	118.43	118.41
C	60	200	66.99	41.06
D	60	200	58.76	15.53
A	60	300	23.27	18.24
B	60	300	70.32	46.89
C	60	300	56.92	20.80
D	60	300	53.15	23.31
A	60	732	82.59	44.32
B	100	732	81.79	53.70
C	150	732	41.48	10.33
D	200	732	43.78	26.50

Figure 11

5 CONCLUSION

In this project we experimented with applying time series forecasting models with different input methods to conduct empirical evaluations of their prediction accuracy. The result indicated the related factors influencing the model's performances and directions of enhancing the model's performance by tuning the hyper-parameters. The future works considering the LSTM itself will include the correlations between multiple companies, and making predictions on multiple attributes.

On the other hand, the Graph Attention Network (GAN) uses temporal attention to make models more accurate. A price encoder and sentiment encoder work to calculate the encoding scheme. Stock price movement, along with sentiment analysis on tweets, or is used to be classify the up or down movement of nodes. The Universal Sentence Encoder (USE) model is used to encode tweet information in a vector. this information is then used to perform sentiment analysis. The analyzed tweets are then inputted into a feedforward neural network. This, along with the graph, is used to predict the up or down movement of nodes in order to predict stock price. Our future work will include the implementation of both the GAN and text encoder; and make further comparisons with the time-series forecasting models.

6 REFERENCES

Weiwei Jiang (2020). Applications of deep learning in stock market prediction: recent progress. Tingwei, Yueting (2021). Improving Stock Closing Price Prediction Using Recurrent Neural Network and Technical Indicators

David M. Q. Nelson, Adriano C. M. Pereira, Renato A. de Oliveira. (2017). Stock market's price movement prediction with lstm neural networks

Alexander, Konstantinos, Micheal, Valentin. (2019).GluonTS: Probabilistic Time Series Models in Python

Konstantinos, Syama, Valentin, Bernie, Danielle, Caner, Jan, Michael, David, Lorenzo, Laurent,Tim (2020). Neural forecasting: Introduction and literature overview Ian, Jorge (2016). Principal component analysis: a review and recent developments

Tingwei, Yueting (2021). Improving Stock Closing Price Prediction Using Recurrent Neural Network and Technical Indicators