Dean&Francis
ISSN 2959-6157

Exploring the Impact of Neural Network Architectures on Prediction Accuracy in Complex Datasets

Minghuen Tai*

School of Professional Studies, New York University, New York, North American

*Corresponding author: mht316@nyu.edu

Abstract:

In this work, to forecast the stock market prices a deep-learning neural network has been employed. The standard statistical models that have been primarily used up until now appear to be inaccurate and time-consuming, and most importantly, they have failed to accurately predict the complex behavior of the data. There are different neural networks present, each with its pros and cons and applications; however, to get the desired results recurrent neural networks (RNNs) are chosen to forecast the stock prices. RNN when used in its classical form also shows some limitations, it has a problem of vanishing gradient, which affects the handling of data. Researchers proposed a modified version of this architecture which is called Long Short-Term Memory (LSTM) architecture. The financial dataset of Apple Inc. used in this work is downloaded from Yahoo Finance. To gauge the working of the model in forecasting Apple Inc.'s stock prices several metrics are used. The metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R2).

Keywords: Stock Market, Neural Networks, RNN, LSTM.

1. Introduction

In the current day and age, artificial intelligence (AI) is one of the hot topics specifically related to research. This phenomenon influences nearly every domain and aspect of professional advancement. Large data models are in use now in abundance to automate the process which usually takes time and needs hours and hours of work. In recent years, financial markets have also used AI and machine learning (ML) to predict stock market prices. Financial markets adopted this new technology because of the complex nature of the data present in the analysis [1]. The forecasting of stock market trends is a widely studied subject in the realms of finance and data analysis. Both financiers and investigators aim to achieve returns from their assets by forecasting the movements of the broader market or specific stocks. Nonetheless, surpassing experienced and educated rivals in the stock markets is a difficult endeavour. Accurate forecasting of stock price movements is essential for effective decision-making and the development of

ISSN 2959-6157

trading strategies. Additionally, it provides shareholders with advance notice regarding unexpected declines in the stock market, particularly concerning immediate expenditures. The standard statistical models that have been primarily used up until now appear to be inaccurate and time-consuming, and most importantly, they have failed to accurately predict the complex behavior of the data. Several factors influence stock prices, including the current market situation, economic data, and geopolitical conditions. These factors make the stock price data difficult to analyze and predict.

Traditional techniques used in forecasting the stock price include autoregressive integrated moving averages (ARIMA) and exponential smoothing approaches [2]. These techniques demonstrated success in prediction when working on simpler patterns and less complex data; however, when employed on complex data, they are less accurate and don't show the characteristics of the data precisely. The advancements in AI have introduced several techniques, mainly neural network structures that have demonstrated strong capability in accurately presenting complex patterns in sequential data. When utilizing neural networks for stock estimation, an operational challenge emerges. The stock exchange is intricate and diverse, readily influenced by economic and political elements. One more challenge in forecasting stock prices with artificial neural networks is selecting a sufficient number of samples, which is based on actual trade information from a specific timeframe. A limited sample size indicates a brief duration of record keeping, which may not be adequate for developing a predictive model. Conversely, an extensive sample size can heighten the unpredictability of the economy throughout the study's duration. The growth of a nation's stock exchange is strongly linked to its economic progress. When there are substantial shifts in a nation's economy, the stock in the marketplace can expect extreme volatility, rendering previous experiences insufficient for forecasting its future movements. There are different neural networks present, each with its pros and cons and applications; however, to get the desired results RNNs are chosen to forecast stock prices built on the data [3].

RNNs have cyclical connections that preserve information from inputs and work robustly on sequential data. The RNN's nature makes it suitable for predicting patterns in financial time series, where past data and figures play an important role. RNN when used in its classical form also shows some limitations, it has a problem of vanishing gradient [4], which affects the handling of data with long-term dependencies. This problem also affects the model's training when trained on a big chunk of sequential data [5]. To address the issue of classical RNN researchers proposed a modified version of this architecture which is called LSTM architecture [6], the modified version mitigates the inherent problems in the RNN. LSTM includes a memory cell in its framework which addresses the problem of vanishing gradient and can learn over a larger chunk of sequential data. Furthermore, LSTM also includes gates in its structure which control the transmission of information through layers and select specific information to be retained or dropped during the training process [7]. This characteristic of LSTM makes it the go-to approach for the training and forecasting of sequential forms of data. The complex nature of the financial data and the workings of the LSTM architecture make it popular as a preferred network to predict stock prices.

The study also leverages the capabilities of LSTM architecture to forecast the data related to stock prices. The choice of network is based on LSTM's ability to handle sequential data modeling better. The study uses stock data from Apple Inc. (AAPL). Financial time series data such as those gathered from Apple Inc., which includes stock prices, frequently display intricate patterns of figures, encompassing trends that carried over a longer period, vulnerability, and abrupt volatility spikes of the data [8]. Predictive models must therefore not only identify immediate pricing fluctuations but also more general, prospective market trends that may have an impact on the organization's future market behavior.

The study aims to investigate the effectiveness of neural network networks, specifically LSTMs, which are modified versions of classical RNNs, in enhancing the prediction accuracy of stock price forecasting. The dataset for this work has been obtained from Yahoo Finance. The dataset of Apple Inc. provides a complete perspective on Apple's stock performance from 2010 to 2023, including different market conditions such as growth rate, lack of progress, and instability. To effectively train and evaluate the model and accurately predict stock market prices a comprehensive dataset is needed. MSE, RMSE, MAE and R2 metrics have been utilized to evaluate the performance of the model [9]. The model's capacity to forecast stock prices consistently is evaluated using an evaluation method that also serves to calculate the model's error rate.

2. Methodology

2.1 Dataset Selection

The financial dataset of Apple Inc. used in this work is downloaded from Yahoo Finance. Several reasons motivated us to choose Apple Inc.'s financial data, which include the international market actively trading Apple's stock. Apple data displays different market behaviors throughout the years, which makes it a suitable selection for predicting financial values utilizing a deep learning network. The dataset covers financial figures from January 2010 to January 2023, covering a variety of marketplaces and periods, including high development phases, market reinforcement corrections, and market volatility. All these characteristics and conditions are important for building a strong predictive model that can work well in varied markets and data. The dataset contains figures of daily stock price data, which includes the Open, High, Low, Close, Adjusted Close, and Volume variables in the.csv file. In this study, the close price has been chosen as the target variable for prediction because it represents the final price at the end of a full trading day and is always considered an important measure in market analysis. Before training the data, the data underwent preprocessing by applying the MinMaxScaler method using Python to normalize the stock prices, scaling them to a range of 0 to 1. Normalizing the data is important in optimizing the neural network's performance. Normalization facilitates robust and faster convergence, preventing the model's learning from being biased towards a specific range of values in the data [10].

2.2 Dataset Pre-Processing

When dealing with time series forecasting of the data, the association between current and past stock prices is important. Time-lagged features are incorporated to improve the time-dependence feature of the LSTM model. The time-lagged feature will allow the model to forecast the stock price of the coming day from the financial data of the preceding 60 days. The sliding-window operation added to the model converts the time series data of the stocks into a format that can be learned by the model. The input feature matrix $X(t)$ and the desired target output $y(t)$ are defined as:

$$X(t) = [x(t-60), x(t-59), \ldots, x(t-1)]$$
$$y(t) = x(t) \tag{1}$$

Where the variable $X(t)$ denotes the stock prices observed throughout the preceding 60 days in the data, whereas $y(t)$ corresponds to the stock price observed on the t-th day, which is the next day of the time series in the financial data. The LSTM used in this work not only integrates the price fluctuations of the current day but also takes into consideration historical trends that are crucial for making financial predictions in the long term. Before training the dataset is partitioned into training and testing sets, allocating 80% of the data for model training and the remaining 20% for testing purposes. The training set is developed to facilitate the model's learning of knowledge from past patterns and figures, while the test set data is developed to assess the model's capacity to generalize to new data and evaluation. Considering the sequential nature of the financial data used in this work, a simple random split might disturb the sequential order. Therefore, an adjacent split was used to maintain the linear sequence of the stock price data of Apple Inc.

2.3 RNN with LSTM

A modified version of RNNs has been implemented in this work, the LSTM architecture. In this work, the modified architecture serves as the fundamental model. CNN is known for focusing on the majority of prominent aspects within the visual field, making it a popular choice in feature creation. LSTM is known for its ability to grow concerning the order of time, making it a popular choice for historical analysis. LSTM network is implemented to address the issue of disappearing gradients, an inherent problem in RNNs. The modified LSTM architecture contains gates that determine the information to retain from the previous step. The gate selects the information from the previous time step to disregard. The gate can be defined as:

$$f_t = \sigma\left(O_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

Where f_t is the vector representing the forget gate, O_f represents the matrix of weights, h_{t-1} demonstrate the hidden state from the previous time step, x_t is the input at the current time step of the data, and b_f is the bias of the system. A sigmoid activation function (σ) is also implemented at the output of the forget gate layer to generate values ranging from 0 to 1, which represent the ratio of information to be kept or discarded during training. The function of the input gate is to regulate and execute the integration of new information into the cell state. The CNN LSTM architecture and LSTM memory cell is demonstrated in Fig. 1, and Fig. 2 respectively. The occurrence can be characterized using the subsequent equations:

$$i_t = \sigma\left(O_i \cdot [h_{t-1}, x_t] + b_i\right)$$
$$\tilde{C}_t = tanh\left(O_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3}$$

Where i_t is the input gate vector, and $\tilde{C}_t$ demonstrate the candidate cell state of the network that captures new information from the data, regulated by the weights O_i and O_C respectively. The final decision on what to output or target from the LSTM gate cell is made by the output gate, which can be defined as:

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right)$$
$$h_t = o_t \cdot tanh\left(C_t\right) \tag{4}$$

ISSN 2959-6157

Where o_t represents the output gate, and h_t represent the hidden state that carries information to the next time step and layer and contributes to the model's prediction capability.

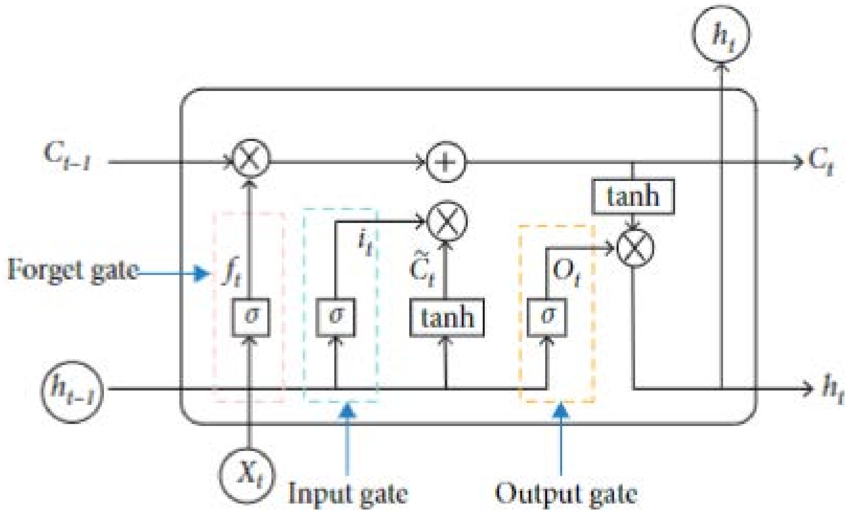

Fig. 1 CNN LSTM Architecture [11]

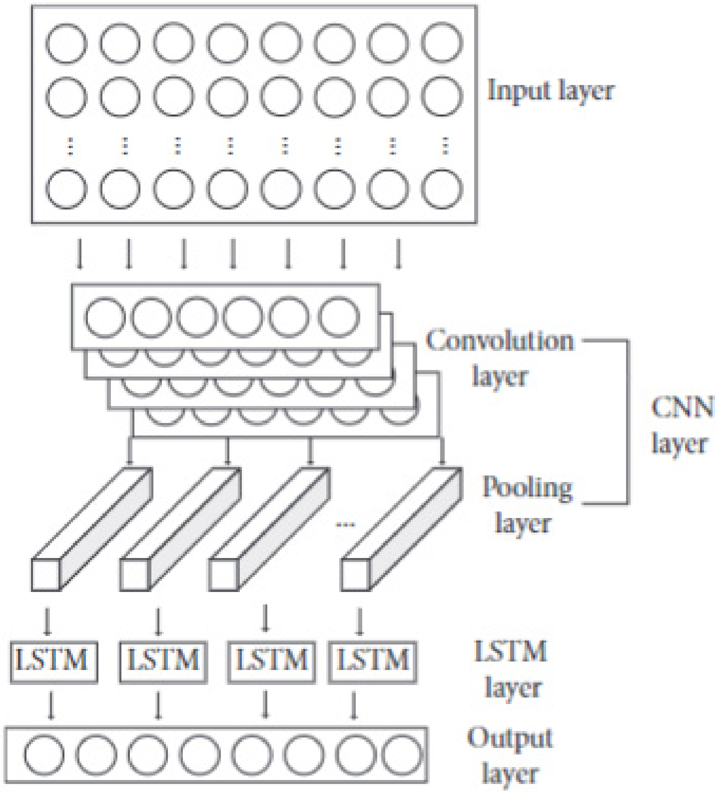

Fig. 2 LSTM memory cell architecture [11]

2.4 Results

To gauge the performance of the LSTM model in predicting Apple Inc.'s stock prices several metrics are used. The metrics include MSE, RMSE, MAE, and R2. These metrics help in quantifying the results of the LSTM model. The results of these metrics are presented in Table 1.

Table 1. Metrics Performance on the Test Set

Metric	Value
MSE	24.7133
RMSE	4.9712
MAE	3.9468
R^2	0.9438

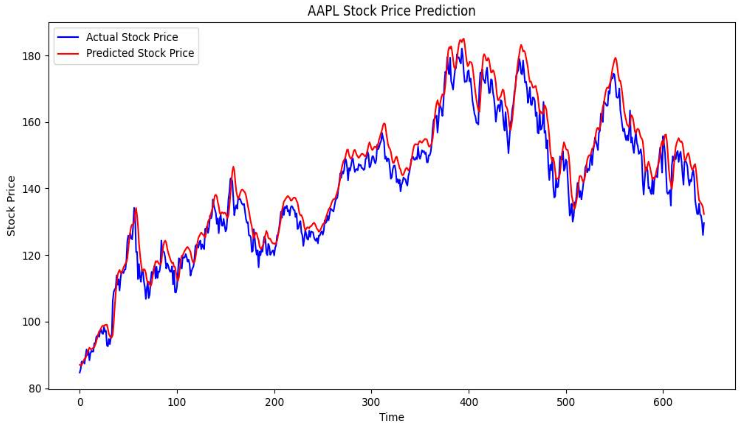

Fig. 3 AAPL Stock Price Prediction Graph

Fig. 3 demonstrates the trend of the prediction over time, the actual stock price is shown in the blue color line, while the predicted stock price is shown in the red color line.

3. Conclusion

This work presents the working of a deep-learning neural network to forecast stock market values. Most notably, the conventional statistical models that have been the primary form of analysis up to this point tend to be laborious, imprecise, and unable to forecast the intricate behaviour of the data. While there are many neural networks available, each with advantages and disadvantages as well as uses, RNNs are selected for forecasting stock prices based on data to achieve the intended outcomes. In its traditional form, RNN exhibits certain limitations as well. For example, it struggles with vanishing gradient, which makes managing data with long-term dependencies difficult. The LSTM architecture is a modified version of this architecture that researchers presented to address the problem with classical RNN. The Apple Inc. financial dataset utilised in this study was obtained from Yahoo Finance. Several measures are used to assess how well the LSTM model predicts the stock prices of Apple Inc. Among the measures are R2, MAE, RMSE, and MSE. The results demonstrate that the model successfully predicts the values and obtains comparable results.

References

[1] Mokhtari Sasan, Yen Kai Kit, and Liu Jiansheng. Effectiveness of artificial intelligence in stock market prediction based on machine learning. *International Journal of Computer Applications*, 2021, 183(7): 1-8.

[2] Xiao Daiyou and Su Jinxia. Research on stock price time series prediction based on deep learning and autoregressive integrated moving average. *Scientific Programming*, 2022, 2022(1): 4758698.

[3] Yoshihara Akihiro, Fujikawa Kazuhiro, Seki Kenta, et al. Predicting stock market trends by recurrent deep neural networks. *In*: Pham Dinh-Nguyen and Park Sang-Bum (eds). PRICAI 2014: Trends in Artificial Intelligence. Cham: Springer International Publishing, 2014: 759-769.

[4] Ribeiro Alexandre H, Tiels Koen, Aguirre Luis A, et al. Beyond exploding and vanishing gradients: analysing RNN training using attractors and smoothness. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. PMLR, 2020: 2370-2380. Available at: https://proceedings.mlr.press/v108/ribeiro20a.html.

[5] Wang Jianhong, Hong Shijun, Dong Yali, et al. Predicting stock market trends using LSTM networks: overcoming RNN limitations for improved financial forecasting. *Journal of Computer Science and Software Applications*, 2024, 4(3): 1-7.

[6] Kumar Subramani Suresh, D Chandra, and Rajan Suresh. Stock price prediction using deep learning LSTM (long short-

ISSN 2959-6157

term memory). *In*: Proceedings of the 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, 2022: 1787-1791. Available at: https://ieeexplore.ieee.org/abstract/document/9823639.
[7] Jenckel Marc, Bukhari Syed Saqib, and Dengel Andreas. Training LSTM-RNN with imperfect transcription: limitations and outcomes. *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing*. New York, NY, USA: Association for Computing Machinery, 2017: 48-53. Available at: https://doi.org/10.1145/3151509.3151527.
[8] Raubitzek Stefan and Neubauer Thomas. An exploratory study on the complexity and machine learning predictability of stock market data. *Entropy*, 2022, 24(3): 332.
[9] Hu Zijing, Zhao Yifan, and Khushi Mohammed. A survey of forex and stock price prediction using deep learning. *Applied System Innovation*, 2021, 4(1): 9.
[10] Pan Jianwei, Zhuang Yiyuan, and Fong Simon. The impact of data normalization on stock market prediction: using SVM and technical indicators. *In*: Berry Michael W, Mohamed Abdul HJ, and Yap Bee Wah (eds). *Soft Computing in Data Science*. Singapore: Springer, 2016: 72-88.
[11] Lu W., Li J., Li Y., et al. A CNN-LSTM-Based Model to Forecast Stock Prices. Journal of Forecasting,[2024] Availablefrom:https://onlinelibrary.wiley.com/doi/10.1155/2020/662297.